

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:12 : Search time 61.04 Seconds

(without alignments)  
331.593 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666

Sequence: 1 WYHYSEKMNQRRARFCR.....NNDACKMNDACKKLAALC 117

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: SPREMBL\_19: \*  
2: sp\_archaea: \*  
3: sp\_bacteria: \*  
4: sp\_fungi: \*  
5: sp\_human: \*  
6: sp\_invertebrate: \*  
7: sp\_mammal: \*  
8: sp\_mhc: \*  
9: sp\_organelle: \*  
10: sp\_phage: \*  
11: sp\_plant: \*  
12: sp\_rhodent: \*  
13: sp\_virus: \*  
14: sp\_vertebrate: \*  
15: sp\_unclassified: \*  
16: sp\_virus: \*  
17: sp\_bacteria: \*  
18: sp\_archaea: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	385	4 Q9UJ43	Q9UJ43 homo sapien
2	598	89.8	376	6 Q28629	Q28629 oryctolagus
3	557	83.6	372	11 Q63762	Q63762 rattus norv
4	445	66.8	616	4 Q95509	Q95509 homo sapien
5	445	66.8	740	4 Q95507	Q95507 homo sapien
6	445	66.8	740	4 Q95508	Q95508 homo sapien
7	440	66.1	484	6 Q951G2	Q951G2 ovus aries
8	437	65.6	485	6 Q951G3	Q951G3 odocoileus
9	434	65.2	754	6 Q28290	Q28290 canis faml1
10	425	63.8	482	6 Q28657	Q28657 oryctolagus
11	420	63.1	649	6 Q28657	Q28657 oryctolagus
12	408	61.3	609	6 Q9GLR0	Q9GLR0 canis faml1
13	401	60.2	646	6 Q29097	Q29097 sus scrofa
14	392.5	58.9	610	6 Q951G1	Q951G1 equus caball
15	177	26.6	328	4 Q9UJ71	Q9UJ71 homo sapien
16	166	24.9	359	5 Q9VQX3	Q9VQX3 drosophila

17	162.5	24.4	381	6 Q95LA8	Q95LA8 macaca mula
18	161.5	24.2	268	4 Q96Q03	Q96Q03 homo sapien
19	161.5	24.2	312	4 Q96Q07	Q96Q07 homo sapien
20	161.5	24.2	334	4 Q96Q09	Q96Q09 homo sapien
21	161.5	24.2	360	4 Q96Q04	Q96Q04 homo sapien
22	161.5	24.2	380	4 Q96Q05	Q96Q05 homo sapien
23	161.5	24.2	380	4 Q96Q00	Q96Q00 homo sapien
24	161.5	24.2	381	6 Q95LC6	Q95LC6 macaca neme
25	161.5	24.2	404	4 Q9NNX6	Q9NNX6 homo sapien
26	161.5	24.2	404	4 Q96Q01	Q96Q01 homo sapien
27	160.5	24.1	292	4 Q14538	Q14538 homo sapien
28	160.5	24.1	404	6 Q951L9	Q951L9 pan troglod
29	160.5	24.1	404	6 Q95J96	Q95J96 macaca mula
30	153.5	23.0	304	11 Q91YT3	Q91YT3 mus musculu
31	150.5	22.6	287	4 Q03969	Q03969 homo sapien
32	150.5	22.6	292	4 Q00448	Q00448 homo sapien
33	149.5	22.4	214	11 Q9RQ08	Q9RQ08 mus musculu
34	148	22.2	178	11 Q91ZW9	Q91ZW9 mus musculu
35	148	22.2	208	11 Q91ZW7	Q91ZW7 mus musculu
36	148	22.2	1479	4 Q9UBG0	Q9UBG0 homo sapien
37	148	22.2	1479	4 Q9Y5P9	Q9Y5P9 homo sapien
38	148	22.2	1479	11 Q64449	Q64449 mus musculu
39	146.5	22.0	263	4 Q96Q03	Q96Q03 homo sapien
40	146.5	22.0	332	4 Q96Q05	Q96Q05 homo sapien
41	146.5	22.0	376	4 Q9BX53	Q9BX53 homo sapien
42	146.5	22.0	398	4 Q96Q08	Q96Q08 homo sapien
43	146.5	22.0	399	4 Q9H2X3	Q9H2X3 homo sapien
44	146	21.9	295	11 Q91ZW4	Q91ZW4 mus musculu
45	146	21.9	311	11 Q9DBV4	Q9DBV4 mus musculu

## ALIGNMENTS

RESULT 1  
Q9UJ43 PRELIMINARY; PRT; 385 AA.  
AC Q9UJ43:  
DT 01-MAY-2000 (TREMBLrel. 13. Created)  
DT 01-MAY-2000 (TREMBLrel. 13. Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19. Last annotation update)  
DE L-SELECTIN PRECURSOR.  
GN L-SELECTIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-HEMATOPOIETIC (B LYMPHOCYTE);  
RA Fieger C.B.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
[2]  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-HEMATOPOIETIC (B LYMPHOCYTE);  
RA Fieger C.B.;  
RL Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.  
DR EMBL: AJ246000; CAB5488.1; -.  
DR HSSP: P14151; IKB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR01304; lectin\_c.  
DR InterPro: IPR02396; selectin.  
DR InterPro: IPR00436; Sush1\_SCR\_CCP.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00059; lectin\_c; 1.  
DR Pfam: PF00084; sush1; 2.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP; 2.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS00615; C\_type\_LECTIN\_1; 1.  
DR PROSITE: PS00441; C\_type\_LECTIN\_2; 1.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.

DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; glycoprotein; Lectin; Selectin; Signal.  
FT SIGNAL 1 51 POTENTIAL.  
FT CHAIN 52 385 L-SELECTIN.  
SQ SEQUENCE 385 AA; 43617 MW; 1205F691BA638EF1 CRC64;

Query Match 100.0%; Score 666; DB 4; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.8e-59;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 60  
DB 52 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 111  
OY 61 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 117  
DB 112 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 168

## RESULT 2

Q28629 PRELIMINARY; PRT; 376 AA.  
AC Q28629;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE L-SELECTIN PRECURSOR.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY CORTEX;  
RA Qian J., Marks R.M.;  
RT "cDNA for rabbit L-selectin."  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U26535; AAA67896.1; -  
DR HSSP; P14151; 1KJB.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001304; Lectin\_c.  
DR InterPro; IPR002396; Selectin.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00008; EGF\_1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi\_2.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00343; CCP; 2.  
DR SMART; SM00344; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; glycoprotein; Lectin; Selectin; Signal.  
FT SIGNAL 1 38 POTENTIAL.  
FT CHAIN 39 376 L-SELECTIN.  
SQ SEQUENCE 376 AA; 42346 MW; 59F6AD530F490947 CRC64;

Query Match 89.8%; Score 598; DB 6; Length 376;  
Best Local Similarity 88.0%; Pred. No. 1.3e-52;  
Matches 103; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 60  
DB 39 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 98  
OY 61 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 117  
DB 99 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 155

RESULT 3  
O63762 PRELIMINARY; PRT; 372 AA.  
AC O63762;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE LYMPHOCYTE MEMBRANE PROTEIN A.11.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=95369821; Pubmed-7543874;  
RA Sackstein R., Meng L., Xu X.M., Chin Y.H.;  
RT "Evidence of post-transcriptional regulation of L-selectin gene  
expression in rat lymphoid cells."  
RL Immunology 85:198-204(1995).  
DR EMBL; S79523; AAC60710.2; -  
DR HSSP; P14151; 1KJB.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001304; Lectin\_c.  
DR InterPro; IPR002396; Selectin.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR Pfam; PF00008; EGF\_1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi\_2.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00343; CCP; 2.  
DR SMART; SM00344; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; glycoprotein.  
SQ SEQUENCE 372 AA; 42471 MW; 7DFD125610DD664A CRC64;

Query Match 83.6%; Score 557; DB 11; Length 372;  
Best Local Similarity 81.2%; Pred. No. 1.9e-48;  
Matches 95; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 60  
DB 39 WYHYSEKPMNWMORARFCNDNTDVAIONKAEIYEKTLPEFSRSYWGIRKIGIW 98  
OY 61 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 117  
DB 99 TWGTTNKSILTEEAENNGDGEPPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLAALC 155

## RESULT 4

O95509 PRELIMINARY; PRT; 616 AA.  
AC O95509;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
DE D1780M13.1.3 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN  
CD62, GMP140)) (ISOFORM 3) (FRAGMENT).  
GN SELP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howden P.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

```
DR EMBL: AL022146; CAA18144.1; -.
DR HSSP: P16109; IKJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_6.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
DR EGF-like domain; glycoprotein.
KW EGF-like domain; glycoprotein.
FT NON_TER 616
SQ SEQUENCE 616 AA; 67736 MW; 35CD4BFADE61D724 CRC64;

Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 616;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKALEYLEKTLPSRSYYWIGIRKIGIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWVGTKSLTDEAENWGDEPNKKKEDCYEITYIKRKNDAGKWNDDACHRLKALC 117
DB 102 TWVGTKKALTNEAENWADNEPNKRNEDCYEITYIKSPSAPGKNWDECHLKKHALC 158

RESULT 5
OY 095507 PRELIMINARY; PRT; 740 AA.
AC 095507;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE DJ780M13.1.1 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN
DE CD62, GMP140)) (ISOFORM 1) (FRAGMENT).
GN SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022146; CAA18142.1; -.
DR HSSP: P16109; IKJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein.
FT NON_TER 740
SQ SEQUENCE 740 AA; 81095 MW; 385F70A45B1A3CD4 CRC64;
```

```
Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 740;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKALEYLEKTLPSRSYYWIGIRKIGIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWVGTKSLTDEAENWGDEPNKKKEDCYEITYIKRKNDAGKWNDDACHRLKALC 117
DB 102 TWVGTKKALTNEAENWADNEPNKRNEDCYEITYIKSPSAPGKNWDECHLKKHALC 158

RESULT 6
OY 095508 PRELIMINARY; PRT; 740 AA.
AC 095508;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)
DE DJ780M13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN
DE CD62, GMP140)) (ISOFORM 2) (FRAGMENT).
GN SELP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Howden P.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL022146; CAA18143.1; -.
DR HSSP: P16109; IKJD.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sush1_8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF_1.
DR PROSITE: PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
DR PROSITE: PS01186; EGF_2; 1.
KW EGF-like domain; glycoprotein.
FT NON_TER 740
SQ SEQUENCE 740 AA; 81390 MW; ID2E35E6D93745CE CRC64;

Query Match
Best Local Similarity 66.8%; Score 445; DB 4; Length 740;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMWQRRARRCRDNYTDVAIQNKALEYLEKTLPSRSYYWIGIRKIGIW 60
DB 42 WTYHSTAYKSWNISRKQCNRYTDVAIQNKNEDYLNKVLPIYSSTYYWIGIRKNNKTW 101
OY 61 TWVGTKSLTDEAENWGDEPNKKKEDCYEITYIKRKNDAGKWNDDACHRLKALC 117
DB 102 TWVGTKKALTNEAENWADNEPNKRNEDCYEITYIKSPSAPGKNWDECHLKKHALC 158

RESULT 7
OY 0951G2 PRELIMINARY; PRT; 484 AA.
AC 0951G2;
DT 01-DEC-2001 (TREMblrel. 19, Created)
```

DT 01-DEC-2001 (TEMBLrel. 19, last sequence update)  
 DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)  
 DE E-SELECTIN.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 CX NCBI\_TaxID=9940;  
 RN [1]  
 RP  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21421234; PubMed=11529941;  
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
 RA MacLachlan N.J.;  
 RT "Characterization of equine E-selectin."  
 RL Immunology 103:498-504(2001).  
 DR EMBL; AF307971; AAK48711.1; -.  
 KW Lectin; Selectin.  
 SQ SEQUENCE 484 AA; 52951 MW; 617930C1C2F47B44 CRC64;

Query Match	66.1%;	Score 440;	DB 6;	Length 484;
Best Local Similarity	66.7%;	Pred. No. 1.9e-36;		
Matches 78;	Conservative 11;	Mismatches 28;	Indels 0;	Gaps 0;

```
QY      1 WTTHXSEKPMNQRRARRRCRDNTDVLAIQNKAIEYLEKTLPFSRSYYWIGIRRGIM    60  
        |::| ::||:: |||||||::| |::::::::::|  
Db     22 WSYHASTEIMFEARDYCCKTYALVALIQOEETELNSTFSPSYWIGIRINGTW    81
```

```

Qy      61 TWVGTSKSLTEAEWMDGDEPPNNKKNKEDCEYIYIKRNKDAGKANDDACHIKKALC 117
      ||:|||||:| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      82 TWIGTNSLTEATWAPGEPNNKQTDDECEVEIYIKREKDSGKANDENCITKKFALC 138

```

RESULT	8			
095L63				
ID	095L63	PRELIMINARY;	PRT;	485 AA.
AC	095L63;			
DT	01-DEC-2001	(TREMBIrel, 19, Created)		
DT	01-DEC-2001	(TREMBIrel, 19, Last sequence update)		
DT	01-DEC-2001	(TREMBIrel, 19, Last annotation update)		
DE	E-SELECTIN.			
OS	Odocolleus hemionus (Mule deer) (Black-tailed deer).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervoidae			
OC	Cervidae; Odocolleinae; Odocolleus.			
OX	NCBI_TaxID=9872;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21421234; Pubmed=11529941;			
RA	Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,			
RA	McLaughlin N.J.;			
RT	"Characterization of equine E-selectin."			
RL	Immunology 103:498-504(2001).			
DR	EMBL; AF307970; AAK48710.1; -.			
DR	Leclat; Selectin.			
Q0	SEQUENCE 485 AA; 53247 MW; 699595199EAFE9980 CRC64;			

Query Match	65.6%;	Score 437;	DB 6;	Length 485;
Best Local Similarity	65.0%;	Pred. No. 3.8e-36;		
Matches	76;	Conservative	15;	Mismatches 26;
			Indels	0;
			Gaps	0;

Qy 61 TWYGTAKSLTEAEWNGDEPPNKKKNKEDCEYETIKRNNKAGCWNDADACHKLALC 117  
Db 83 TWYGTAKSLTKATWAPDEPNKNSODDCEYETIKREKSGGWNDEBCTKRRLAC 139

RESULT 9  
Q28290

ID	Q28290	PRELIMINARY;	PRT;	754 AA.
AC	Q28290.			
DT	01-NOV-1996	(TRMBIrel, 01, Created)		
DT	01-NOV-1996	(TRMBIrel, 01, Last sequence update)		
DT	01-DEC-2001	(TRMBIrel, 19, Last annotation update)		
DE	CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).			
GN	GMP140.			
OS	Canis familiaris (Dog).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.			
OX	NCBI_TaxID=9615;			

RP SEQUENCE FROM N.A.  
RC TISSUE-SPLEEN;  
RA Manning A.M., Sanders W.E.Jr. , Kujielka G.L., Dore M.,  
RA Rosenbloom C.L., Hawkins H.L., Michael L.H., Smith C.W.,  
RA Beutner A.L., Anderson D.C.;  
RT "Molecular cloning of canine GMP140 and studies of expression in a  
RT model of myocardial ischemia/reperfusion",  
RL Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.

DR	InterPro: IPR000561; EGF-like.		
DR	InterPro: IPR001304; lectin.C		
DR	InterPro: IPR002396; selectin.		
DR	InterPro: IPR000436; Sushi_SCR_CCP.		
DR	Pfam: PF00059; lectin.c; 1.		
DR	Pfam: PF00084; sushi; 8.		
DR	PRINTS: PR00343; SELECTIN.		
DR	SMART: SM00032; CCP; 8.		
DR	SMART: SM00034; CLECT; 1.		
DR	SMART: SM00181; EGF; 1.		
DR	PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.		
DR	PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.		
DR	PROSITE: PS00022; EGF; 1; UNKNOWN_1.		
DR	PROSITE: PS01186; EGF_2; 1.		
KW	EGF-like domain; Glycoprotein; signal.		
FT	NON_PEP	1	
FT	SIGNAL	<1	13
EQ	CHAIN	14	82303
EQ	SEQUENCE	754 AA;	POTENTIAL. CELL ADHESION MOLECULE. F0438BEAA521E773 CRC64;

Query Match	65.28;	Score 434;	DB 6;	Length 754;
Best Local Similarity	65.08;	Pred. No. 1.3e-35;		
Matches 76;	Conservative 10;	Mismatches 31;	Indels 0;	Gaps 0;

QY 1 WTHYSKPKPMNRRARFCRDYATDVLAIQKAAEEYEKTLSPRSRWYIGRTGGTW 60  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 27 WTYNSFKRAISWYSKIFQGHYTDVAIQKKELAYINDYIPYNSYATGIRKINDMY 86  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 QY 61 TWVGTGKSLTEEFENMGDSEPPNNKKNKDCVEYIYTKRRKDAGKNDCAKRLKALIC 117  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :  
 Db 87 TWVGTKKPLTEEFENMAEENPPNNKNNDQCEVIYIKSLAPAGKNDSEPPCKRRKRAL 143  
 ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :

RESULT 10

ID	Q28982	PRELIMINARY;	PRT;	482 AA.
----	--------	--------------	------	---------

DT	01-NOV-1996	(TREMblurel, 01, Created)
DT	01-NOV-1996	(TREMblurel, 01, last sequence update)
DT	01-DEC-2001	(TREMblurel, 19, last annotation update)

OS *Sus scrofa* (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
OX NCBI\_TaxID=9823;

RA Bach F.H.;  
RA Winkler H., Brostjan C., Csizmadia V., Natarajan G., Anrather J.,  
RA MEDLINE=97075911; Pubmed=8918234;  
RX CORRELATION WITH



RT "The Intron-exon structure of the porcine E-selectin-encoding gene."  
 RL Gene 176:67-72(1996).  
 DR EMBL: U37521; AAC48680.1; -.  
 DR HSSP: P16581; IESL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; Lectin\_C\_1.  
 DR Pfam: PF00084; Sush1\_4.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP\_4.  
 DR SMART: SM00034; CLECT\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM EGF-like domain; Glycoprotein; Lectin; Selectin.  
 SO SEQUENCE 482 AA; 52341 MW; 97DCSD70BF115944 CRC64;

Query Match 63.8%; Score 425; DB 6; Length 482;  
 Best Local Similarity 63.2%; Pred. No. 6.1e-35;  
 Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMWQRRARFCRDNYTDLVAIQNKAETLEYEKLTPESRSYWIWIGIRKGIW 60  
 DB 23 WSYNSTEAMTFDEASTYCCOQRYTHLVAIQNKAETLEYINSTFNFSASYWIGIRKNGTW 82  
 OY 61 TWVGNTKSLTEAEWNGDEPNKKKEDCVETIKRKNDGKMWDDACHLKAALC 117  
 DB 83 TWIGTKALTPEDATNMAPGEPPNNKQNDCEVEITIKRKDGSKMWDERCKRKALC 139

RESULT 11  
 ID Q28657 PRELIMINARY: PRT: 649 AA.  
 AC Q28657;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE P-SELECTIN.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vora D.K., Fang Z., Liva S.M., Parhami F., Watson A.D., Drake T.A.,  
 RA Territo M.C., Berliner J.A.;  
 RT "Induction of p-selectin by MW-LDL and its role in human  
 RT atherosclerosis."  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Warden C.H.;  
 RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U39446; AAA81385.1; -.  
 DR HSSP: P16109; IFSB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; Lectin\_C\_1.  
 DR Pfam: PF00084; Sush1\_6.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP\_6.  
 DR SMART: SM00034; CLECT\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 KM Lectin; Selectin.  
 SO SEQUENCE 609 AA; 66073 MW; 41E62D1FAD23881F CRC64;

DR PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KM EGF-like domain; Glycoprotein; Lectin; Selectin.  
 SO SEQUENCE 649 AA; 71755 MW; ECCD8C847B84BC31 CRC64;

Query Match 63.1%; Score 420; DB 6; Length 649;  
 Best Local Similarity 65.0%; Pred. No. 2.8e-34;  
 Matches 78; Conservative 11; Mismatches 25; Indels 6; Gaps 3;

OY 1 WTYHYSEKPMWQRRARFCRDNYTDLVAIQNKAETLEYEKLTPESRSYWIWIGIRK----IG 57  
 DB 42 WTYHYSNKTYSNVNRACQKRYTDLVAIQNKNEDLYNETIPYNSYWIWIGIRKDDQIN 101  
 OY 58 GIWTVGNTKSLTEAEWNGDEPNKKKEDCVETIKRKNDGKMWDDACHLKAALC 117  
 DB 102 G--TWVG-NKNTLEAEWNGDEPNKKKEDCVETIKRSLAPGKWDEPCWKRRKALC 158

RESULT 12

ID Q9GLF0 PRELIMINARY: PRT: 609 AA.  
 AC Q9GLF0;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Canis familiaris (Dog).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Carnivora; Fissipedata; Canidae; Canis.  
 OX NCBI\_TaxID=9615;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-ENDOTHELIAL CELLS;  
 RA Zheng L., Shi Y., Wu H., Zhang G.;  
 RT "Cloning and sequencing of beagle E-selectin genomic DNA and  
 RT comparison with other species."  
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF287257; AAG10039.1; -.  
 DR HSSP: P16581; IESL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001304; Lectin\_C.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00059; Lectin\_C\_1.  
 DR Pfam: PF00084; Sush1\_6.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP\_6.  
 DR SMART: SM00034; CLECT\_1.  
 DR SMART: SM00181; EGF\_4.  
 DR SMART: SM00001; EGF-like; 1.  
 DR PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 KM Lectin; Selectin.  
 SO SEQUENCE 609 AA; 66073 MW; 41E62D1FAD23881F CRC64;

Query Match 61.3%; Score 408; DB 6; Length 609;  
 Best Local Similarity 59.0%; Pred. No. 4.3e-33;  
 Matches 69; Conservative 18; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMWQRRARFCRDNYTDLVAIQNKAETLEYEKLTPESRSYWIWIGIRKGIW 60  
 DB 23 WSYNSTEAMTFDEASTYCCOQRYTHLVAIQNKAETLEYINSTFNFSASYWIGIRKNGTW 82  
 OY 61 TWVGNTKSLTEAEWNGDEPNKKKEDCVETIKRKNDGKMWDDACHLKAALC 117  
 DB 83 TWIGTKALTPEDATNMAPGEPPNNKQNDCEVEITIKRKDGSKMWDERCKRKALC 139

RESULT 13  
 Q29097

```

ID 029097 PRELIMINARY; PRT; 646 AA.
AC 029097;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE P-SELECTIN PRECURSOR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=AORTA.
RA Rollins S.A., Johnson K.K., Birks C.W., Matis L.A., Rother R.P.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=AORTA.
RA MEDLINE=20171534; Pubmed=10706724;
RX Stocker C.J., Sugars K.L., Harari O.A., Landis R.C., Morley B.J.,
RA Haskard D.O.;
RT "TNF-alpha, IL-4, and IFN-gamma regulate differential expression of P-
RT and E-selectin expression by porcine aortic endothelial cells.";
RL J. Immunol. 164:3309-3315(2000).
DR EMBL; L39075; AAR79007.1; -.
DR EMBL; AF163766; AAF43272.1; -.
DR HSSP; P16109; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001304; Lectin_C.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; Lectin_C; 1.
DR Pfam; PF00084; Sushi; 6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.
DR PROSITE; PS01186; EGF_2; 1.
DR EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.
FT SIGNAL 1 41
FT CHAIN 42 646 P-SELECTIN.
FT SEQUENCE 646 AA; 71127 MW; 3863FAFE09F0DB6 CRC64;

Query Match 60.2%; Score 401; DB 6; Length 646;
Best Local Similarity 59.0%; Pred. No. 2,4e-32;
Matches 69; Conservative 14; Mismatches 34; Indels 0; Gaps 0;

OY 1 WYHYSEKPMNMOARPCRDNTDVAIONKAEIEYLEKTLPPSRSYWIGIRKIGIW 60
DB 42 WYNYSTSAVSNMTSRVFCQRYFTDLVAIONKKEIAYINDVIRYSSYWIGIRKINNK 101
OY 61 TWVGTKSLTEEAENMGDEPNKKKEDCVETIYIKRNKDAGKWDACHKLAALC 117
DB 102 TWVGTKSLTEEAENMGDEPNKKKEDCVETIYIKRNKDAGKWDACHKLAALC 158

RESULT 14
O95LGI PRELIMINARY; PRT; 610 AA.
AC 095LGI;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE E-SELECTIN.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
NCBI_TaxID=9796;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21421234; Pubmed=11529941;
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.L.,
RA MacLachlan N.J.;
RT "Characterization of equine E-selectin.";
RL Immunology 103:498-504(2001).
DR EMBL; AF307972; AAK48712.1; -.
KW Lectin; Selectin.
SQ SEQUENCE 610 AA; 66191 MW; F9D3DED12C445382 CRC64;

Query Match 58.9%; Score 392.5; DB 6; Length 610;
Best Local Similarity 60.7%; Pred. No. 1,6e-31;
Matches 71; Conservative 12; Mismatches 33; Indels 1; Gaps 1;

OY 1 WYHYSEKPMNMOARPCRDNTDVAIONKAEIEYLEKTLPPSRSYWIGIRKIGIW 60
DB 22 WSYASATTNMTFDEASAYCOORYTHLVAIONBEIKYLNISIFNHSPSYWIGIRKYNCKW 81
OY 61 TWVGTKSLTEEAENMGDEPNKKKEDCVETIYIKRNKDAGKWDACHKLAALC 117
DB 82 TWVGTKSLTEEAENMGDEPNKKKEDCVETIYIKRNKDAGKWDACHKLAALC 137

RESULT 15
O9U71 PRELIMINARY; PRT; 328 AA.
AC 09U71;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE LANGERIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Valladeau J., Ravel O., Dezutter-Dambuyant C., Moore K., Kleijmeer M.,
RA Duvallet-Frances V., Vincent C., Schmitt D., Davoust J., Caux C.,
RA Lebecque S., Saeland S.;
RT "Langerin, a new transmembrane C-type lectin specific to langerhans
RT cells, induces the formation of Birbeck granules.";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ242859; CAB62403.1; -.
DR HSSP; P20693; IHLJ.
DR InterPro: IPR001304; Lectin_C.
DR Pfam; PF00059; Lectin_C; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
SQ SEQUENCE 328 AA; 36697 MW; 83DF5A1B347D1B62 CRC64;

Query Match 26.6%; Score 177; DB 4; Length 328;
Best Local Similarity 34.7%; Pred. No. 5,3e-10;
Matches 41; Conservative 16; Mismatches 53; Indels 8; Gaps 4;

OY 3 YHYSEKPMNMOARPCRDNTDVAIONKAEIEYLEKTLPPSRSYWIGIRKIGIW 60
DB 207 YFSLIPKTIWYSAEFCVSRNSHLTSVTSSEDEFLYKTA--GGLIYWIGLKAMEGDW 264
OY 61 TWVGTKSLTEEAENMGDEPNKKKEDCVETIYIKRNKDAGKWDACHKLAALC 117
DB 265 SWVDITPFKVKYSAFWIIPDEPNAGNNEHCNT---KAPSLQAMNDAPCDKTLPLTFC 319

RESULT 16
O9VOX3 PRELIMINARY; PRT; 359 AA.
AC 09VOX3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)

```

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE CG3958 PROTEIN (H105328P).  
 CN LECTIN-24DB OR CG2958.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NCBI\_TaxID=7227;  
 RN  
 RP  
 SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Fertler S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jallat M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Klamme B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclob J.M.,  
 RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Stryckas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RT Science 287:2185-2195(2000).  
 RL  
 RN  
 RP  
 SEQUENCE FROM N.A.  
 RC STRAIN=Y, CN BW SP;  
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
 RA Champe M., Chavez C., Dorsett V., Fartan D., Frise E., George R.,  
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,  
 RA Nuno J., Paclob J., Paragas V., Park S., Phouanavong S., Wan K.,  
 RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AEO03577; AFS1038.1; -;  
 DR EMBL: AY061025; AAL28573.1; -;  
 DR HSSP: P14151; 1KJB  
 DR Piybase: FBgn0040102; lectin-24Db.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR SMART: SM00034; CLECT; 1.  
 DR PROSITE: PSS0041; C-TYPE LECTIN.2; 1.  
 SQ SEQUENCE 359 AA; 41546 MW; 1895E525594584C4 CRC64;

Query Match 24.9%; Score 166; DB 5; Length 359;  
 Best Local Similarity 33.6%; Pred. No. 7.6e-09;

Matches 36; Conservative 16; Mismatches 47; Indels 8; Gaps 2;  
 QY 11 NMQRARRCRDNYTDLVAIONKAELEYEKLTPFSRSYWGIRKIGGWTWVGKSLT 70  
 DB 255 DWQSAVDFCRWGGYIAIKDDELDIAISARL--DQKSYWGINDLQSSNNYVSVASGRE 312  
 QY 71 EAEWNGGDEPNKKKEDCEIYIKRNKDKAGWDDACHKKAALC 117  
 DB 313 VEFLLMNMNGEPHNGEDNCVELT-----RSKMNDDPCHRRKHVYC 353  
 RESULT 17  
 0951AB PRELIMINARY; PRT; 381 AA.  
 AC 0951AB:  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE DENDRITIC CELL-SPECIFIC ICAM-3 GRABING NONINTEGRIN.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 NCBI\_TaxID=9544;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21465051; PubMed=11581396;  
 RA Baribud F., Pohlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,  
 RA Haggarty B.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,  
 RA Arneson J., Reinhardt T.A., Kimata J.T., Littman D.R., Hoxie J.A.,  
 RA Doms R.W.;  
 RT "Functional and antigenic characterization of human, rhesus macaque,  
 RT pitallied macaque, and murine dc-sign.";  
 RL J. Virol. 75:10281-10289(2001).  
 DR EMBL: AF369755; AAL14438.1; -;  
 KW Integrin.  
 SQ SEQUENCE 381 AA; 42897 MW; 01E7B0B42C91D49 CRC64;  
 Query Match 24.4%; Score 162.5; DB 6; Length 381;  
 Best Local Similarity 29.6%; Pred. No. 1.9e-08;  
 Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4;  
 QY 1 WT-----YHSEKPMNQRRARPCRDNYTDLVAIONKAELEYEKLTPFSRSYWGIR 54  
 DB 237 WTFQGNCFYFNSORNNHNSITACQEGADLVYKSAEONFLOLQSSRNRFWGLS 296  
 QY 55 KIG--GIWTWGTNKSLEAEADNMGDEPNKKKEDCEIYIKRNKDKAGWDDACHKL 112  
 DB 297 DLNHEGTQWWDGSLPLPSFQYNNKGEPPN--IGBEDCAEP-----SGNGMNDKCNLA 349  
 QY 113 KAALC 117  
 DB 350 KFWIC 354  
 RESULT 18  
 0960Q3 PRELIMINARY; PRT; 268 AA.  
 AC 0960Q3:  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE SDC-STGNIA TYPE III ISOFORM.  
 GN CD209.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX PubMed=11337487;  
 SQ Mummidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,

RA Jimenez F., Ahuja S.S., Ahuja S.K.;  
RT "Extensive Repertoire of Membrane-bound and Soluble Pentameric Cell-  
RT specific ICM-3-grabbing NonIntegrin I (DC-SIGN) and DC-SIGN2  
RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN  
RT Transcripts.";  
RT J. Biol. Chem. 276:33196-33212(2001).  
DR EMBL: AY042227; AAK91852.1; -  
SO SEQUENCE 268 AA: 30427 MW: 113E1ED1B5748693 CRC64;

Query Match	24.28;	Score 161.5;	DB 4;	Length 268;
Best Local Similarity	29.68;	Pred. No. 1.5e-08;		
Matches 37;	Conservative 16;	Mismatches 57;	Indels 15;	Gaps 4

```

QY      1 WT-----YHSEKPMWQRRRCGNDYDVAIQKAELEYLEKTLPPSRSTYWTGIR 54
Db      124 WTFQGCACYFMSNORMWHDISTACKEVGAQLVYIKSAEONFQIOLOSSRNRETFWGLS 183
QY      55 KIG--GIWTWGTNKSLTLEAENWGDEPPNNKKNECDCEVELYIKRNDACKWDDACHKL 112
Db      184 DLNDEGTWQWYDSSPLLPFRQYINRGEPPNN--VGEEDCAEF-----SGNWMDDKCNLA 236
QY      113 KAAKL 117
Db      237 KFWIC 241

```

RESULT	19		
096007			
ID	096007	PRELIMINARY;	PRT; 312 AA.
AC	096007;		
DT	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)		
DE	MDC-SIGNIA TYPE III ISOFORM.		
CN	CD209.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	PubMed=11337487;		
RA	Mummidli S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,		
RA	Jimenez F., Ahuja S.S., Ahuja S.K.;		
RT	"Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-		
RT	specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2		
RT	Isoforms. Inter-individual Variation in Expression of DC-SIGN		
RT	Transcripts. "		
RL	J. Biol. Chem. 276:33196-33212(2001).		
DR	EMBL: AY042223; AAK91848.1; -		
SO	SEQUENCE 312 AA; 3517 MM; 95761618EF36AFBE CRC64;		

```

Query Match          24.2%; Score 161.5; DB 4; Length 312;
Best Local Similarity 29.6%; Pred. No. 1.8e-08;
Matches 37; Conservative 16; Mismatches 57; Indels 15; Gaps 4.

QY 1 WT-----YHSEKEMNNQRRARFCRDYTDVAIONKAEIYELEKTLPEFSRSYYWIGIR 54
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 168 WTFEGGNCYCPMNSNQRNHHDDSTTACKKEVGAOLVYIKSAEQNFLOQSSRSNRFYTWGILS 227
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 55 KIG--GIWTVGTSKLTLEAEENWGDGEENNNKKNFEDCEVEIYIKRNKDGAKWNDDACHKL 112
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 228 DLNDEGTQWQWVDGSPLLPSFKQYNNRGEPNN-VGEEDCAEF-----SGGNNDKCNLA 280
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

QY 113 KAALC 117
    ||  ||
Db 281 KFWIC 285

```

RESULT 20  
Q96QP9

ID	Q960P9	PRELIMINARY;	PR1;	334 AA.
AC	Q960P9;			
DT	01-DEC-2001 (TREMBLrel, 19, 'Created')			
DT	01-DEC-2001 (TREMBLrel, 19, last sequence update)			
DT	01-DEC-2001 (TREMBLrel, 19, last annotation update)			
DE	SDC-SIGN1B TYPE II ISOFORM.			
GN	CD209.			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=11337487;			
RA	Mummidli S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,			
RA	Jimenez F., Ahuja S.S., Ahuja S.K.;			
RT	"Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-			
RT	specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2			
RT	Isoforms. Inter-Individual Variation In Expression of DC-SIGN			
RT	Transcripts.";			
RL	J. Biol. Chem. 276:33196-33212(2001).			
DR	EMBL; AY042231; AAK91856.1; -			
SO	SEQUENCE 334 AA; 37843 MW; 8E796FC4111C86BF CRC64;			

Query Match	24.2%;	Score 161.5;	DB 4;	Length 334;
Best Local Similarity	29.6%;	Pred. No. 2e-08;		
Matches 37; Conservative	16;	Mismatches 57;	Indels 15;	Gaps 4;

OY 1 WT-----THYESEKPYNNQARAFECRDHYTDVLAIOKAKEIELEKTLPFSYSYWGIR 54  
Db 190 WTEFGNCFEENSQQRNMHDSITACKELGVALYIKSAEQGNLTLOSSRSRFTPMGLS 24.9

OY 55 KIG--GIIMWYTSTNKSLTLEEANNGDPEPNKKRKEDCVETIYIKRKADGAKNDACHTL 112  
Db 250 DLDQEGTMQWVGSGPLLSFESFOYNRRGPNN--VGEECAEF-----SGNGWNDDCNLA 302

OY 113 KAALC 117  
Db 303 KEWIC 307

RESULT	21	
096004		
ID	096004	PRELIMINARY; PRT; 360 AA.
AC	096004;	
DT	01-DEC-2001 (TrEMBLrel. 19, Created)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)	
DE	SDC-SIGNIA TYPE II ISOFORM.	
GN	CD209.	
OS	Homo sapiens (Human).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eulhelia; Primates; Catarrhini; Homiidae; Homo.	
OX	NCBI_TaxID=9606;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RX	PubMed11337487;	
RA	Mummidli S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,	
RA	Jimenez F., Ahuja S.S., Ahuja S.K.;	
RT	Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-	
RT	specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2	
RT	Isoforms. Inter-Individual Variation in Expression of DC-SIGN	
RT	Transcripts. "	
RL	J. Biol. Chem. 276:33196-33212(2001).	
DR	EMBL: AY042226; AAK91851.1; -	
SQ	SEQUENCE 360 AA; 41009 MM; 6ABE2B9AAEDFAA8 CRC64;	

Query Match	24.2%;	Score 161.5;	DB 4;	Length 360;
Best Local Similarity	29.6%;	Pred. No. 2.2e-08;		
Matches 37;	Conservative 16;	Mismatches 57;	Indels 15;	Gaps 4;

```

OY      1 WT-----YHSEKRMNMNRARPCRDWNTDLVALONKAELIETKLTPEFSRSYWGIR 54
Db      216 WFFEOGNCIFENSNSCRNMHDSITACKEGAGQLVVIKSABEONFLOQSRSRNRFYWGIS 275
OY      55 KIG--GIWTWVTGNKSLTLEAEWMGCDGEPENNKNEDECEYIYIKRNKDAGKNNDDACHLT 112
Db      276 DLNGLCTQWWDGSPILPSPFKQYWRNGEPNN-VGEDCAEF-----SGGNNDCKCNLA 328
OY      113 KAAAC 117
Db      329 KFWIC 333

RESULT  22
O96005
ID      O96005      PRELIMINARY;      PRT;      380 AA.
AC      O96005;
DT      01-DEC-2001 (TREMBLrel. 19, Created)
DT      01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      SDC-SIGNIA TYPE I ISOFORM.
GN      CD209.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      PubMed=11337487;
RA      Mummid S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,
RA      Jimenez F., Ahuja S.S., Ahuja S.K.;
RT      "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-
RT      specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN) and DC-SIGN2
RT      Isoforms. Inter-Individual Variation in Expression of DC-SIGN
RT      Transcripts.", 276:33196-33212(2001).
RL      J. Biol. Chem. 276:33196-33212(2001).
SQ      EMBL; AY042225; AAK91850.1; -.
SQ      SEQUENCE 380 AA; 43330 MW; BCF9CC45ABEF6B02 CRC64;

```

Query Match	24.2%	Score 161.5	DB 4	Length 380
Best Local Similarity	29.6%	Pred. No. 2.3e-08		
Matches 37	Conservative 16	Mismatches 57	Indels 15	Gaps 4
QY	1 WT-----YHYEKKPMNMQRARRFCRDNYTDLVAIQNKAELIYLEKTLFFSRSYWIGIR	54		
Db	236 WFFPGNGCYFMSNSQRNMHDSITACKVEAGQLVVIKSAEQNPLDQSSNRFTFMGLS	295		
QY	55 KTG--GIWTWGTNKSLETEAEWNGDGEFNKKNKEDCEYIITKRKKDGKNNDDACHL	112		
Db	236 DLNQGTMQWVDGSPILPSFKQYNRNGEPNN-VGEEDCAEF-----SGNGNNDKCNLA	348		
QY	113 KALAC 117			
Db	349 KFWIC 353			
RESULT 23				
Q96000				
ID	096000	PRELIMINARY	PRT	380 AA.
AC	096000			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	SDC-SIGNIB TYPE I ISOFORM.			
GN	CD209.			
OS	homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_taxid=9606;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RX	pubmed=11337487;			

RA Mummidge S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,  
 RA Jimenez F., Ahuja S.S., Ahuja S.R.,  
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-  
 RT specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2  
 RT Isoforms. Inter-individual Variation in Expression of DC-SIGN  
 RT Transcripts.", 276:33196-33212(2001).  
 RL J. Biol. Chem.,  
 DL EMBL, AY042230; AAK91855.1; -  
 SO SEQUENCE 380 AA; 43125 MW; F3D098F9FB7D044B CRC64;

Query Match	24.28;	Score 161.5;	DB 4;	Length 380;
Best Local Similarity	29.68;	Pred. No. 2.3e-08;		
Matches 37;	Conservative 16;	Mismatches 57;	Indels 15;	Gaps 4;

```

QY 1 WT-----YHSEKPMNQARRCRDNYDTLVAIQKAELEYLEKTLPPSRSTYWGIR 54
Db 236 WTFEOGNCYEFMSNSQRNMHDSITACKELVGAQLVYIKSAEENFNFIQLOSSRSNRETYMGLS 295
QY 55 KIG--GIYTWGVTGKSLLEEDAMNGDGBEPNNKKKKEDCEVLEYIKRNKDAGWMDACHTL 112
Db 296 DLNMGSTQWVWDGSPLLPSFKQYNNRGPNN-VGEEDCAEF-----SGNGWMDDKCNLA 348
QY 113 KAALC 117
Db 349 KFWIC 353

```

RESULT	24			
Q95LC6	ID	Q95LC6	PRELIMINARY;	PRT; 381 AA.
AC	Q95LC6;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	DENDRITIC CELL-SPECIFIC ICAM-3 GRABING NONINTEGRIN.			
OS	Macaca nemestrina (Pig-tailed macaque).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;			
OC	Cercopitheciinae; Macaca.			
NCBI_TaxID=9545;				
RM	[1]			
RP	SEQUENCE FROM N.A.			
FX	MEDLINE=21465051; PubMed=11581396;			
RA	Baribud F., Pohlmann S., Sparwasser T., Kimata M.T., Choi Y.K.,			
RA	Hagagaty B.S., Ahmad N., Macfarlan T., Edwards T.G., Leslie G.J.,			
RA	Aranson J., Reinhart T.A., Kimata J.T., Littman D.R., Hoxle J.A.,			
RA	Doms R.W.;			
RT	"Functional and antigenic characterization of human, rhesus macaque			
RT	pigtailed macaque, and murine dc-sigm.";			
RL	J. Virol. 75:10281-10289(2001).			
DR	EMBL: AF343727; MML14428.1; -.			
KM	Integrin.			
SO	SEQUENCE	381 AA;	42951 MW;	C4F6E2E3D454B7A4 CRC64;

Query Match	24.2%;	Score 161.5;	DB 6;	Length 381;
Best Local Similarity	29.6%;	Pred. No. 2.3e-08;		
Matches 37;	Conservative 16;	Mismatches 57;	Indels. 15;	Gaps 4

```

0Y 1 WT-----YHSEKEMNQARRREFRDNYTLDAVIAONNAEELEYLEKTLPSFRSYVYWGIR 54
Db 237 WTFEGNGCYEPMNSNGRNRHDSITACQEVGAQLVYIKSAEEQNFIQLDSSRSNRPFTWGLS 296
0Y 55 KIG--GIMTWGVTGKSLTEEAENNGDSEBPNKKKKEPCVLEYIRKNDAGKWMNDACHTL 112
Db 297 DLNHGHTQMWADGSEFLPSFKQYNNKGEPNN-VGEEDCAEF-----SGNGMNDKCNLA 349
0Y 113 RAALC 117
Db 350 KFWIC 354

```

ID	Q9NNX6	PRELIMINARY:	PRT:	404 AA.
AC	Q9NNX6			
DT	01-OCT-2000 (TREMBLrel. 15, Created)			
DT	01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	MEMBRANE-ASSOCIATED LECTIN TYPE C (PROBABLE MANNOSE-BINDING C-TYPE LECTIN DC-SIGN) (MOC-SIGNIA TYPE I ISOPROB.)			
CN	CD209.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=PLACENTA;			
RX	MEDLINE=92390446; PubMed=1518869;			
RA	Curtis B.M., Schiarowski S., Watson A.J.;			
RT	"Sequence and expression of a membrane-associated C-type lectin that			
RT	exhibits CD4-independent binding of human immunodeficiency virus			
RT	envelope glycoprotein gp 120.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 89:8356-8360(1992).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20432267; PubMed=10975799;			
RA	Solileux E.J., Barten R., Trowsdale J.;			
RT	"DC-SIGN, a related gene, DC-SIGNR, and CD23 form a cluster on			
RT	19p13.";			
RL	J. Immunol. 165:2937-2942(2000).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21157496; PubMed=11257134;			
RA	Bashtrova A.A., Geijtenbeek T.B.H., van Duljnhoven G.C.F.,			
RA	van Vliet S.J., Eilering J.B.G., Martin M.P., Wu L., Martin T.D.,			
RA	Vieblis N., Knolle P.A., KewalRamani V.N., van Kooyk Y., Carrington M.			
RT	"A dendritic cell-specific intercellular adhesion molecule 3-grabbling			
RT	nonintegrin (DC-SIGN)-related protein is highly expressed on human			
RT	liver sinusoidal endothelial cells and promotes HIV-1 infection.";			
RL	J. Exp. Med. 193:671-671(2001).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	Pubmed=11337487.			
RA	Mumidi S., Catano G., Lam L., Hoefle A., Telles V., Begum K.,			
RT	Jimenez F., Ahuja S.S., Ahuja S.K.;			
RT	"Extensive Repertoire of Membrane-Bound and Soluble Dendritic Cell-			
RT	specific ICM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2			
RT	isoforms. Inter-individual Variation in Expression of DC-SIGN			
RT	Transcripts.";			
RL	J. Biol. Chem. 276:33196-33212(2001).			
DR	EMBL: M98457; AAF77072.1; -			
DR	EMBL: AF209479; AAG13814.1; -			
DR	EMBL: AF290886; AAK20997.1; -			
DR	EMBL: AY042221; AAK91846.1; -			
DR	HSSP: P20693; 1HLJ.			
DR	InterPro: IPR001304; lectin_c.			
DR	Pfam: PF00059; lectin_c; 1.			
DR	SMART: SM00034; CLECT; 1.			
DR	PROSITE: PS00615; C_TYPE_LECTIN_1; UNKNOWN_1.			
DR	PROSITE: PS50041; C_TYPE_LECTIN_2; 1.			
DR	lectin.			
SO	SEQUENCE 404 AA; 45775 MW; A23FA24601453C0 CRC64;			
QY	Query Match 24.28; Score 161.5; DB 4; Length 404;			
Db	Best Local Similarity 29.66; Pred. No. 2.5e-08;			
Matches	37; Conservative 16; Mismatches 57; Indels 15; Gaps			
1	WT-----YHSEKPMNQBARRCNDNYDVAIONKAEIETLEKTLPSRSYWNIGR 54			
260	WTFPGNCYFNSQNRNHDSTIRACKEGVGAOLYVTKSAEDONFLOLOSSNRFWTWGLS 319			
55	KIG--GIMTWGINKSLTEAENMGDEPNKKKNEDECVETIYIKRNKDACRWNDACCHK 1122			

[illegible]



```

RESULT 31
003969
ID 003969 PRELIMINARY; PRT; 287 AA.
AC 003969;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HEPATIC LECTIN H2, VARIANT (ASIALOGLYCOPROTEIN RECEPTOR H2).
GN L-H2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=92184202; PubMed=1371982;
RA Paletta E., Racevskis J., Stockert R.J.;
RT "Differences in the abundance of variably spliced transcripts for the
RT second asialoglycoprotein receptor polypeptide, H2, in normal and
RT transformed human liver.";
RL Hepatology 15:395-402(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=HEPATOMA;
RA Strausberg R.;
RT Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBUNIT: THE FUNCTIONING LIGAND-BINDING UNIT OF THIS RECEPTOR IS
CC THOUGHT TO BE AT LEAST A DIMER.
CC -1- MISCELLANEOUS: THIS PROTEIN IS A VARIANT OF HUMAN HEPATIC LECTIN
CC H2 (AC P07307).
CC -1- MISCELLANEOUS: THE ASIALOGLYCOPROTEIN RECEPTOR IS EXPRESSED
CC EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS IN MAMMALS. AFTER LIGAND
CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND
CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE
CC DIASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE
CC SURFACE.
CC -1- MISCELLANEOUS: AFTER REMOVAL OF STATIC ACID MONOMERS FROM THE
CC COMPLEX CARBOHYDRATE MOETIES OF PLASMA GLYCOPROTEINS, TERMINAL
CC GALACTOSE UNITS ARE RECOGNIZED BY THIS HEPATIC RECEPTOR.
CC -1- MISCELLANEOUS: THE UNUSUAL ORIENTATION OF THIS PROTEIN ACROSS THE
CC MEMBRANE IS POSTULATED TO OCCUR BY RECOGNITION OF AN INTERNAL
CC TRANSMEMBRANE SEGMENT AS A SIGNAL SEQUENCE.
DR EMBL: X55283; CAA38997.1; -.
DR EMBL: BC017251; AAH17251.1; -.
DR HSSP; P06734; 1HTI.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Lectin, glycoprotein, Receptor, Endocytosis, Transmembrane.
FT CAROHD 78 78 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CAROHD 146 146 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT CAROHD 281 281 N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ SEQUENCE 287 AA; 32550 MW; E9528ACFCB1435 CRC64;

Query Match 22.6%; Score 150.5; DB 4; Length 287;
Best Local Similarity 30.3%; Pred. No. 2.2e-07;
Matches 37; Conservative 21; Mismatches 47; Indels 17; Gaps 6;

```

```

Db 275 VC 276
!
RESULT 32
000448
ID 000448 PRELIMINARY; PRT; 292 AA.
AC 000448;
DT 01-JUL-1997 (TREMBlrel. 04, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ASIALOGLYCOPROTEIN RECEPTOR.
GN ASGPR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=86016723; PubMed=3663106;
RA Spiess M., Lodish H.F.;
RT "Sequence of a second human asialoglycoprotein receptor: conservation
RT of two receptor genes during evolution.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:6465-6469(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92184202; PubMed=1371982;
RA Paletta E., Stockert R.J., Racevskis J.;
RT "Differences in the abundance of variably spliced transcripts for the
RT second asialoglycoprotein receptor polypeptide, H2, in normal and
RT transformed human liver.";
RL Hepatology 15:395-402(1992).
RN [3]
RP SEQUENCE FROM N.A.
RA Lee D.G., Lee S.G., Kim K.L., Hahn K.-S.;
RT "Sequences of cDNAs encoding two subunits of asialoglycoprotein
RT receptor in human liver.";
RL submitted (APR-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: U97197; AAB58308.1; -.
DR HSSP; P06734; 1HTI.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF000059; lectin_c; 1.
DR SMART; SM00034; CLECT; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
KW Receptor.
SQ SEQUENCE 292 AA; 33087 MW; 8DC85A2B9757CA9A CRC64;

Query Match 22.6%; Score 150.5; DB 4; Length 292;
Best Local Similarity 30.3%; Pred. No. 2.2e-07;
Matches 37; Conservative 21; Mismatches 47; Indels 17; Gaps 6;

```





```

RA Behrendt N., Jensen O.N., Engelholm L.H., Mortz E., Mann M., Dano K.;
RT "A urokinase receptor-associated protein with specific collagen-
RT binding properties.";
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=98403880; PubMed=9734811;
RA Ishikawa K., Nagase T., Suyama M., Miyajima N., Tanaka A., Kotani H.,
RA Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. X.
RT The complete sequences of 100 new cDNA clones from brain which can
RT code for large proteins in vitro.";
RL DNA Res. 5:169-176(1998).
DR EMBL; AF107292; AAF1492.1; -.
DR EMBL; AB014609; BAA31684.1; -.
DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; lectin_c.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE; PS50213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Receptor; Kinase.
SQ SEQUENCE 1479 AA; 16654 MW; C7583EA78E2792D1 CRC64;

Query Match 22.2%; Score 148; DB 4; Length 1479;
Best Local Similarity 30.3%; Pred. No. 2.8e-06;
Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHSEKPMNQRRRCRDNYTDLVAIONKAELEYLEKTLPSRSYWGIR--KIGGIW 60
DB 394 YRLQAKRRSWQSKKACLRGGDLVSIHSMALERTTKQIKQVEBELWGLNDLKQNMF 453
OY 61 TWV-TGNKSLTEAEAWMGDEPNN-KNKEDCEVEIYIKRNKAGKNNDDACHKLKALC 117
DB 454 EWSDGSLVST----HWHPPEPNFRDLSLEDCVTIW----GPEGRNDSPCNQLSPSIC 504

RESULT 37
OY5P9 PRELIMINARY; PRT; 1479 AA.
AC OY5P9;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ENDOCYTIC RECEPTOR ENDO180.
GN ENDO180.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20148849; PubMed=10683150;
RA Sheikh H., Yarwood H., Ashworth A., Isacke C.;
RT "Endo180, an endocytic recycling glycoprotein related to the
RT macrophage mannose receptor is expressed on fibroblasts, endothelial
RT cells and macrophages and functions as a lectin receptor.";
RL J. Cell Sci. 113:1021-1032(2000).
EMBL; AF134838; AAD30280.1; -.

```

```

DR HSSP; P02751; 2FN2.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; lectin_c.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR Pfam; PF00652; Ricin_B_lectin; 1.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.
DR PROSITE; PS00023; FIBRONECTIN_2; UNKNOWN_1.
DR PROSITE; PS50213; LIPOCALIN; UNKNOWN_1.
DR PROSITE; PS50231; RICIN_B_LECTIN; 1.
KW Receptor.
SQ SEQUENCE 1479 AA; 16669 MW; 9F4BAF355F036FCE CRC64;

Query Match 22.2%; Score 148; DB 4; Length 1479;
Best Local Similarity 30.3%; Pred. No. 2.8e-06;
Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHSEKPMNQRRRCRDNYTDLVAIONKAELEYLEKTLPSRSYWGIR--KIGGIW 60
DB 394 YRLQAKRRSWQSKKACLRGGDLVSIHSMALERTTKQIKQVEBELWGLNDLKQNMF 453
OY 61 TWV-TGNKSLTEAEAWMGDEPNN-KNKEDCEVEIYIKRNKAGKNNDDACHKLKALC 117
DB 454 EWSDGSLVST----HWHPPEPNFRDLSLEDCVTIW----GPEGRNDSPCNQLSPSIC 504

RESULT 38
OY6449 PRELIMINARY; PRT; 1479 AA.
AC OY6449;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE LECTIN LAMBDA.
GN MRC2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96355501; PubMed=8702911;
RA Wu K., Yuan J., Lasky L.A.;
RT "Characterization of a novel member of the macrophage mannose receptor
RT type C lectin family.";
RL J. Biol. Chem. 271:21323-21330(1996).
DR EMBL; U56734; AAC52729.1; -.
DR HSSP; P02751; 2FN2.
DR MGD; MGI:107818; Mrc2.
DR InterPro; IPR001128; Cyt_P450.
DR InterPro; IPR000562; FN_Type_II.
DR InterPro; IPR001304; lectin_c.
DR InterPro; IPR000566; Lipocln_cytfabp.
DR InterPro; IPR000772; Ricin_B_lectin.
DR Pfam; PF00040; fn2; 1.
DR Pfam; PF00059; lectin_c; 8.
DR PRINTS; PR00013; FNTYPEII.
DR ProDom; PD000995; FN_Type_II; 1.
DR SMART; SM00034; CLECT; 8.
DR SMART; SM00059; FN2; 1.
DR SMART; SM00458; RICIN; 1.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 3.
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 8.

```

DR PROSITE: PS00023; FIBRONECTIN\_2; UNKNOWN\_1.  
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN\_1.  
 DR PROSITE: PS50231; RICIN B LECTIN; 1.  
 SO SEQUENCE 1479 AA; 167112 MW; 620456E109B48C1 CRC64;

Query Match 22.0%; Score 148; DB 11; Length 1479;  
 Best Local Similarity 30.3%; Pred. No. 2.8e-06;  
 Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

OY 3 YHYSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR--NIGSIW 60  
 DB 393 YRLQAEKRSMQESKRACLRGGDLISHSMALEIFITQIKQVEBELWIGLNDLTKQNMFF 452  
 OY 61 TWV-GTNKSLTEEAENMGDGEPPN-KKNEKDCVEIYIKRNDAGKWNDDACHKLKAALC 117  
 DB 453 EMSDGSIVSFT---HMHPEPNRFDLSLDCVITW---GPEGKWNDSPCNQLSPSIC 503

RESULT 39  
 O960P3 PRELIMINARY; PRT: 263 AA.

AC O960P3; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE SDC-SIGN2 TYPE III ISOFORM.  
 CN CD209L.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX PubMed-1137487;  
 RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,  
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;  
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-  
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2  
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN  
 RT Transcripts";  
 RL J. Biol. Chem. 276:33196-33212(2001).  
 DR EMBL: AY042240; AAK91865.1;--  
 SO SEQUENCE 263 AA; 30102 MW; C611C762C50424CF CRC64;

Query Match 22.0%; Score 146.5; DB 4; Length 263;  
 Best Local Similarity 29.1%; Pred. No. 5e-07;  
 Matches 34; Conservative 12; Mismatches 56; Indels 15; Gaps 4;

OY 1 WT-----YHSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR 54  
 DB 136 WTFQGNCFYFMSNSQRNHDVTAQCEVRAQLVYIKTAEQNFLOLQTSRSNRFSGMGLS 195  
 OY 55 KIG--GIWTWGTNKSLEEAENMGDGEPPNKKNEKDCVEIYIKRNDAGKWNDDAC 109  
 DB 196 DLNDEGTQWWDGSPSPSFPQRYWNSGEPNNSGN-EDCAEF-----SGSGWMDNRC 245

RESULT 40  
 O960P5 PRELIMINARY; PRT: 332 AA.  
 AC O960P5; 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE SDC-SIGN2 TYPE I ISOFORM.  
 CN CD209L.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX PubMed-1137487;  
 RA Mummidi S., Catano G., Lam L., Hoeftle A., Telles V., Begum K.,  
 RA Jimenez F., Ahuja S.S., Ahuja S.K.;  
 RT "Extensive Repertoire of Membrane-bound and Soluble Dendritic Cell-  
 RT Specific ICAM-3-grabbing Nonintegrin 1 (DC-SIGN1) and DC-SIGN2  
 RT Isoforms. Inter-Individual Variation in Expression of DC-SIGN  
 RT Transcripts";  
 RL J. Biol. Chem. 276:33196-33212(2001).  
 DR EMBL: AY042238; AAK91863.1;--  
 SO SEQUENCE 332 AA; 37953 MW; CGFDEP92C1B073C6 CRC64;

Query Match 22.0%; Score 146.5; DB 4; Length 332;  
 Best Local Similarity 29.1%; Pred. No. 6.6e-07;  
 Matches 34; Conservative 12; Mismatches 56; Indels 15; Gaps 4;

OY 1 WT-----YHSEKPMNQRRARPCRDNTDLVAIONKAEIYELEKTLPPSRSYWIGIR 54  
 DB 205 WTFQGNCFYFMSNSQRNHDVTAQCEVRAQLVYIKTAEQNFLOLQTSRSNRFSGMGLS 264  
 OY 55 KIG--GIWTWGTNKSLEEAENMGDGEPPNKKNEKDCVEIYIKRNDAGKWNDDAC 109  
 DB 265 DLNDEGTQWWDGSPSPSFPQRYWNSGEPNNSGN-EDCAEF-----SGSGWMDNRC 314

Search completed: September 7, 2002, 10:23:17  
 Job time: 545 sec

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:47 : Search time 18.65 seconds  
(without alignments)  
242.905 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666  
Sequence: 1 WTYHYSEKMNQRRARFCR.....NKDAGKMNDACHKLKALIC 117

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SWISSProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	% Match	Query Length	ID	Description
1	666	100.0	372	1	LEM1_HUMAN
2	666	100.0	372	1	LEM1_PANTR
3	662	99.4	372	1	LEM1_PONPY
4	631	94.7	372	1	LEM1_MACMU
5	631	94.7	372	1	LEM1_PAPHA
6	581	87.2	370	1	LEM1_BOVIN
7	557	84.8	372	1	LEM1_MOUSE
8	557	83.6	372	1	LEM1_RAT
9	464	69.7	769	1	LEM3_SHEEP
10	455	68.3	768	1	LEM3_MOUSE
11	451	67.7	646	1	LEM3_BOVIN
12	445	66.8	485	1	LEM2_BOVIN
13	445	66.8	830	1	LEM3_HUMAN
14	442	66.4	768	1	LEM3_RAT
15	434	65.2	551	1	LEM2_RABIT
16	425	63.8	484	1	LEM2_PIG
17	415	62.3	610	1	LEM2_HUMAN
18	408	61.3	549	1	LEM2_RAT
19	408	61.3	611	1	LEM2_CANFA
20	405	60.8	612	1	LEM2_MOUSE
21	157	23.6	1456	1	MANR_HUMAN
22	154.5	23.2	321	1	FCE2_HUMAN
23	153.5	23.0	304	1	MMGL_MOUSE
24	153	23.0	548	1	KOCR_MOUSE
25	150.5	22.6	311	1	LECI_HUMAN
26	149	22.4	207	1	LECH_CHICK
27	143.5	21.5	301	1	LECI_MOUSE
28	141.5	21.2	301	1	LECI_RAT
29	141.5	21.2	331	1	FCE2_MOUSE
30	133.5	20.0	306	1	MMGL_RAT
31	133	20.0	158	1	LECG_TRIST
32	132	19.8	550	1	KOCR_RAT
33	129.5	19.4	283	1	LECH_MOUSE

34	126	18.9	374	1	PSPD_MOUSE
35	125	18.8	374	1	PSPD_RAT
36	123.5	18.5	244	1	MABC_MOUSE
37	122	18.3	162	1	LECG3_MEGMO
38	121	18.2	165	1	LITH_RAT
39	121	18.2	2124	1	PCGA_RAT
40	121	18.2	2132	1	PCGA_MOUSE
41	120.5	18.1	283	1	LECH_RAT
42	120	18.0	2333	1	PCGA_CANFA
43	119.5	17.9	273	1	MT75_MOUSE
44	118.5	17.8	2415	1	PCGA_HUMAN
45	118	17.7	3562	1	PCGV_CHICK

## ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	P14151	P15023			
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SELL OR LYAM1 OR LNH1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-89315837; PubMed-2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary conservation at tandem cell interaction domains.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-Tonsil;				
RX	MEDLINE-89310350; PubMed-2473156;				
RA	Tedder T.F., Isaacs C.M., Ernst T.J., Demetri G.D., Adler D.A.,				
RT	Disteche C.M.;				
RT	"Isolation and chromosomal localization of cDNAs encoding a novel human lymphocyte cell surface molecule, LAM-1. Homology with the mouse lymphocyte homing receptor and other human adhesion proteins.";				
RL	J. Exp. Med. 170:123-133(1989).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-90044046; PubMed-2509939;				
RA	Camerlini D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.";				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-89308881; PubMed-2663882;				
RA	Bowen B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph node homing receptor.";				
RL	J. Cell Biol. 109:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE-90243637; PubMed-1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RT	"Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";				
RL	J. Biol. Chem. 265:7760-7767(1990).				
RN	[6]				
RP	3D-STRUCTURE MODELING.				

```

RX MEDLINE=96074584; PubMed=7488174;
RA Bajorath J., Aruffo A.;
RT "A template for generation and comparison of three-dimensional
RT selectin models.";
RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC DATABASE: NAME=PROT; NOTE=CD guide CD62L entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd62l.htm".
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M25280; AAC63053.1; -
DR EMBL: X16150; CAA34275.1; -
DR EMBL: X17519; CAB43536.1; -
DR EMBL: X17519; CAB43537.1; ALT_SEQ.
DR EMBL: X16070; CAA34203.1; ALT_SEQ.
DR EMBL: M32414; AAB60700.1; -
DR EMBL: M32406; AAB60700.1; JOINED.
DR EMBL: M32407; AAB60700.1; JOINED.
DR EMBL: M32408; AAB60700.1; JOINED.
DR EMBL: M32409; AAB60700.1; JOINED.
DR EMBL: M32410; AAB60700.1; JOINED.
DR EMBL: M32411; AAB60700.1; JOINED.
DR EMBL: M32412; AAB60700.1; JOINED.
DR EMBL: M32413; AAB60700.1; JOINED.
DR PIR: A33912; A33912.
DR PIR: A34015; A34015.
DR PIR: S06798; S06798.
DR PIR: J10104; J10104.
DR PDB: 1KJB; 03-APR-96.
DR GLYCOSITE: P14151; -.
DR MIM: 153240; -.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; Lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS00641; C-TYPE_LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Repeat; 3D-structure.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRASMEX 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
BY SIMILARITY.

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 37 37 D -> Y (IN REF. 4).
FT CONFLICT 178 178 Y -> H (IN REF. 4).
FT CONFLICT 193 193 F -> L (IN REF. 1 AND 4).
FT CONFLICT 213 213 L -> S (IN REF. 3).
FT CONFLICT 214 214 L -> F (IN REF. 4).
FT CONFLICT 218 220 SFS -> MFM (IN REF. 2).
FT CONFLICT 242 242 G -> E (IN REF. 2).
SQ SEQUENCE 372 AA; 42187 MW; 6EA9918ECA2D3643 CRC64;

Query Match 100.0%; Score 666; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,8e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNWRARPCRDNYTDVAIONKAIEYLEKTLPFSSSYWIGIRKIGTW 60
DB 39 WYHYSEKPMNWRARPCRDNYTDVAIONKAIEYLEKTLPFSSSYWIGIRKIGTW 98
QY 61 TWYGTNKSLLTEAEENWGDGEPNNKKKEDCEVEYTYRNNDAKWDNDACHKIKALC 117
DB 99 TWYGTNKSLLTEAEENWGDGEPNNKKKEDCEVEYTYRNNDAKWDNDACHKIKALC 155

RESULT 2
ID LEM1_PANTR STANDARD; PRT; 372 AA.
AC 095237;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
GN SELL.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Budhan J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC      or send an email to license@lsb-sib.ch).
CC      -----
CC      EMBL: U73728; AAB18248.1; -.
DR      HSSP; P14151; IKJB.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR002396; Selectin.
DR      InterPro; IPR000436; Sush1_SCR_CCP.
DR      InterPro; IPR001304; lectin_c.
DR      Pfam; PF00008; EGF_1.
DR      Pfam; PF00059; lectin_c; 1.
DR      Pfam; PF00084; sush1; 2.
DR      PRINTS; PR00343; SELECTIN.
DR      SMART; SM00032; CCP; 2.
DR      SMART; SM00034; CLECT; 1.
DR      SMART; SM00181; EGF_1.
DR      PROSITE; PS00022; EGF_1; 1.
DR      PROSITE; PS01186; EGF_2; 1.
DR      PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR      PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW      Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW      Selectin; Signal; Sush1; Repeat.
FT      SIGNAL          1      28
FT      PROPEP          29      38
FT      CHAIN           39      372
FT      DOMAIN          39      332
FT      TRANSMEM        333      355
FT      DOMAIN          356      372
FT      DOMAIN          55      155
FT      DOMAIN          156      192
FT      DOMAIN          196      255
FT      DOMAIN          258      317
FT      DISULFID        57      155
FT      DISULFID        128      147
FT      DISULFID        160      171
FT      DISULFID        165      180
FT      DISULFID        182      191
FT      DISULFID        197      241
FT      DISULFID        227      254
FT      DISULFID        259      303
FT      DISULFID        289      316
FT      CARBOHYD        60
FT      CARBOHYD        104      104
FT      CARBOHYD        176      177
FT      CARBOHYD        216      216
FT      CARBOHYD        232      232
FT      CARBOHYD        246      246
FT      CARBOHYD        271      271
SQ      SEQUENCE       372 AA; 42188 MW; 6EA991802AD3643 CRC64;

Query Match      100.0%; Score 666; DB 1; Length 372;
Best Local Similarity 100.0%; Pred. No. 2.8e-61;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  WYHYSEKPMNMQRARFCRDNYTDLVAIONKAELEYEKLTPFSRSYVWIGIRKIGIM 60
DB      39  WYHYSEKPMNMQRARFCRDNYTDLVAIONKAELEYEKLTPFSRSYVWIGIRKIGIM 98
QY      61  TWVGTKSLTEAEENMGDEPNKKKEDCVETIKRNKDAGKANDDCIKLKALALC 117
DB      99  TWVGTKSLTEAEENMGDEPNKKKEDCVETIKRNKDAGKANDDCIKLKALALC 155

RESULT      3
LEML_PONPY  STANDARD:      PRT:      372 AA.
ID      095235;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion
DE      molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE      (LECAM1) (CD62L).

```

```

GN      SELL.
OS      Pongo pygmaeus (Orangutan).
OC      Eukaryota; Metazoa; Chordata; Creniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX      NCBI_TaxID=9600;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA      Tsurushita N.,
RL      Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC      - FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC      OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC      VENULES IN PERIPHERAL LYMPH NODES.
CC      - SUBCELLULAR LOCATION: Type I membrane protein.
CC      - SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC      - SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      - SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      - SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC      or send an email to license@lsb-sib.ch).
CC      -----
CC      EMBL: U73729; AAB18247.1; -.
DR      HSSP; P14151; IKJB.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR002396; Selectin.
DR      InterPro; IPR000436; Sush1_SCR_CCP.
DR      InterPro; IPR001304; lectin_c.
DR      Pfam; PF00008; EGF_1.
DR      Pfam; PF00059; lectin_c; 1.
DR      PRINTS; PR00343; SELECTIN.
DR      SMART; SM00032; CCP; 2.
DR      SMART; SM00034; CLECT; 1.
DR      SMART; SM00181; EGF_1.
DR      PROSITE; PS00022; EGF_1; 1.
DR      PROSITE; PS01186; EGF_2; 1.
DR      PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR      PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW      Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW      Selectin; Signal; Sush1; Repeat.
FT      SIGNAL          1      28
FT      PROPEP          29      38
FT      CHAIN           39      372
FT      DOMAIN          39      332
FT      TRANSMEM        333      355
FT      DOMAIN          356      372
FT      DOMAIN          55      155
FT      DOMAIN          156      192
FT      DOMAIN          196      255
FT      DOMAIN          258      317
FT      DISULFID        57      155
FT      DISULFID        128      147
FT      DISULFID        160      171
FT      DISULFID        165      180
FT      DISULFID        182      191
FT      DISULFID        197      241
FT      DISULFID        227      254
FT      DISULFID        259      303
FT      DISULFID        289      316
FT      CARBOHYD        60      60
FT      CARBOHYD        104      104
FT      CARBOHYD        177      177
FT      CARBOHYD        216      216
FT      CARBOHYD        226      226
FT      CARBOHYD        232      232
FT      CARBOHYD        246      246
FT      CARBOHYD        271      271
SQ      SEQUENCE       372 AA; 42188 MW; 6EA991802AD3643 CRC64;

```





```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U52074; AAB40903.1; -.
DR HSP: P14151; IKB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sushi_SCR_CCP.
DR InterPro: IPR01304; lectin_c.
DR Pfam: PF00008; EGF_1; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00034; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 1.
DR PROSITE: PS00615; C_Type_Lectin_2; 1.
DR PROSITE: PS00615; C_Type_Lectin_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 372
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
FT DISULFID 160 171
FT DISULFID 165 180
FT DISULFID 182 191
FT DISULFID 197 241
FT DISULFID 227 254
FT DISULFID 229 303
FT DISULFID 289 316
FT CARBOHYD 60 316
FT CARBOHYD 104 316
FT CARBOHYD 177 316
FT CARBOHYD 226 316
FT CARBOHYD 232 316
FT CARBOHYD 246 316
FT CARBOHYD 271 316
SO SEQUENCE 372 AA; 42091 MW; 64E7BDD5AC549D69 CRC64;

Query Match 94.7%; Score 631; DB 1; Length 372;
Best local Similarity 94.0%; Pred. No. 1;le-57;
Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYHSEKPMNQRARFCRDNYTDLVAIONKAEYELEKTPFSRSYWGIRKIGIM 60
DB 39 WYHHSSEPMNQRARFCRDNYTDLVAIONKAEYELEKTPFSRSYWGIRKIGIM 98
OY 61 TWVGTSKSLTEEAENWGDEPNKKKEDCEVEIYIKRNKDGKMWDDACHKLAALC 117
DB 99 TWVGTSKSLTEEAENWGDEPNKKKEDCEVEIYIKRNKDGKMWDDACHKLAALC 155

RESULT 6
LEMI_BOVIN
ID LEMI_BOVIN STANDARD: PRT; 370 AA.
AC P98131;

DR 01-FEB-1996 (Rel. 33; Created)
DR 01-FEB-1996 (Rel. 33; Last sequence update)
DR 01-FEB-1996 (Rel. 33; Last annotation update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
DE SEL.
GN Bos taurus (Bovine).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP MEDLINE=92164727; PubMed=1371468;
RA Walcheck B., White M., Kurk S., Kishimoto T.K., Jutila M.A.;
RT "Characterization of the bovine peripheral lymph node homing
RT receptor: a lectin cell adhesion molecule (LECAM).";
RL Eur. J. Immunol. 22:469-476(1992).
RN [2]
RP MEDLINE=94055053; PubMed=7694420;
RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;
RT "Bovine L-selectin: a peripheral lymphocyte homing receptor.";
RL Vet. Immunol. Immunopathol. 37:201-215(1993).
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X62882; CAA44676.1; -.
DR HSP: P14151; IKB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sushi_SCR_CCP.
DR InterPro: IPR01304; lectin_c.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C_Type_Lectin_1; 1.
DR PROSITE: PS00615; C_Type_Lectin_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 370
FT DOMAIN 39 333
FT TRANSMEM 334 354
FT DOMAIN 355 370
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
```

RP SEQUENCE OF 1-360 FROM N.A.  
RX MEDLINE=91169529; PubMed=2004776;  
RA Dowdenko D.J., Diep A., Taylor B.A., Lusis A.J., Lasky L.A.;

[illegible]

```

FT CARBOHYD 226 226 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 278 278 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 288 288 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 308 308 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT CARBOHYD 320 320 N-LINKED (GLCNAC . . .) (POTENTIAL).
FT COMPLECT 32 32 I -> T (IN REF. 4).
SQ SEQUENCE 372 AA; 42288 MW; 4433BDF6E4CB2B78 CRC64;

Query Match      Best Local Similarity    84.8%; Score 565; DB 1; Length 372;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0.

QY 1 WTYHSEKPMNQARRCRDNYTDLVAIQNKAEILEYELEKTLPFSKYWIGIRKIGIM 60
   |||||:::||||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 39 WTYHSEKPMNENARKCKQKVDTDLVAIQNKREILEYLENTLPKSPYYTWIGIRKIKGM 98
   |||||:::||||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

QY 61 TWVGINKSLTEAEHWGDEFPNNKNKKEDCVETIYIKRNKDAGKANDDACIKALC 117
   |||||:::||||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
Db 99 TWVGINKSLTEAEHWGDEFPNNKNKKEDCVETIYIKRDSGSKWMDACHKRKALC 155
   |||||:::||||:|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 8
LEML_RAT LEML_RAT STANDARD: PRT: 372 AA.
ID AC P30836;
AD 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE l-selectin precursor (lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAW-1) (LY-22) (Lymphocyte surface MEL-14 antigen)
DE (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN SELOR LNHR OR LY-22.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP MEDLINE=92329548; PubMed=1378303;
RX Watanabe T., Song Y., Hirayama Y., Tamatani T., Kuida K., Miyasaka M.;
RA "Sequence and expression of a rat CDNA for LECAM-1."
RL Blochum. Biophys. Acta 1131:321-324(1992).
CC CC -1 FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENUES IN PERIPHERAL LYMPH NODES.
CC CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC CC -1 SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
-----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way used or modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@isb-sdb.ch).
-----
CC DR EMBL: D10831, BAA01613.1; -.
DR PIR: S23936; S23936.
DR HSSP: PL4151; IKTB.
DR InterPro: IPRO00561; EGF-like.
DR InterPro: IPRO02396; Selectin.
DR InterPro: IPRO00436; Sushi_SCP_CCP.
DR InterPro: IPRO01304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_C_1.
DR Pfam: PF00084; sushi_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00322; CCP_2.

```

Query Match	Best Local	Similarity	83.6%	Score 557	DB 1	Length 372
Matches	95	Conservative	11	Mismatches	11	Indels
Qy	1	WTYHSEKPMNQARRRCRDNYTDVLAIONKAELEYLEKTLPSRSYTWIGIRIGSIW	60			
Db	39	WYHYSERSMWMENARKCKHNYTDVLAIONKRELEYLEKTLPKNPYTWIGIRIGKTW	98			
Oy	61	TWGYNKSKSTJTEAEKMGDGEPPNKKRDKCEYIYIKRKDKGKNMDDACHRLKALC	117			
Db	99	TWVGINKLTLTKEAEWNGTGEPPNKKSKEDCYEYIYIKRERDSGKNMDDACHRKALC	155			
RESULT	9					
LEM3_SHEEP	ID	LEM3_SHEEP	STANDARD:	PRT:	769 AA.	
AC	P98109					
DT	01-FEB-1996 (Rel. 33, Created)					
DT	01-FEB-1996 (Rel. 33, Last sequence update)					
DT	16-OCT-2001 (Rel. 40, Last annotation update)					
DE	P-selectin precursor (granule membrane protein 140) (GMP-140) (PADGEM)					
DE	(CP62p) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).					
GN	SELP.					
OS	Ovis aries (sheep).					
OC	Eukaryota, Metazoa, Chordata, Vertebrata, Euteleostomi;					
OC	Mammalia, Eutheria, Cetartiodactyla, Ruminantia, Pecora, Bovidae;					
OC	Bovidae, Caprinae, Ovis.					
OX	NCBI_TaxID=9940;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	TISSUE=Heart;					
RA	Burns S.A., Neufeld E.J., Donady J.J.;					
RL	Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.					
CC	-1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS					
CC	TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE					
CC	INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH					
CC	LEUCOCYTES. THE LIGAND RECOGNIZED IS STAYLX-LEWIS X.					

[illegible]

CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).	
CC	-----	
DR	EMBL; M87861; AAA40008.1; -;	
DR	EMBL; M72332; AAA37712.1; -;	
DR	PIR; A42755; A42755.	
DR	HSSP; P16109; IFSB.	
DR	MCD; MG1:98280; SeLP.	
DR	InterPro; IPRO00561; EGF_1like.	
DR	InterPro; IPRO00742; EGF_2.	
DR	InterPro; IPRO02396; Selectin.	
DR	InterPro; IPRO00436; Sushi_SCC_CCP.	
DR	InterPro; IPRO01304; lectin_c.	
DR	Pfam; PF00008; EGF_1.	
DR	Pfam; PF00059; lectin_c; 1.	
DR	Pfam; PF00084; sushi; 8.	
DR	PRINTS; PR00343; SELECTIN.	
DR	SMART; SM00032; CCP; 8.	
DR	SMART; SM00034; CLECT; 1.	
DR	SMART; SM00181; EGF; 1.	
DR	PROSITE; PS00022; EGF_1; 1.	
DR	PROSITE; PS01186; EGF_2; 1.	
DR	PROSITE; PS00615; C-TYPE_LECTIN_1; 1.	
DR	PROSITE; PS50041; C-TYPE_LECTIN_2; 1.	
KW	Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;	
KW	Selectin; Signal; Sushi; Repeat; Lipoprotein; Palmitate.	
FT	SIGNAL	1
FT	CHAIN	42
FT	DOMAIN	42
FT	TRANSMEM	710
FT	DOMAIN	733
FT	DOMAIN	734
FT	DOMAIN	58
FT	DOMAIN	159
FT	DOMAIN	199
FT	DOMAIN	261
FT	DOMAIN	320
FT	DOMAIN	323
FT	DOMAIN	385
FT	DOMAIN	385
FT	DOMAIN	447
FT	DOMAIN	509
FT	DOMAIN	579
FT	DOMAIN	641
FT	DISULFID	60
FT	DISULFID	131
FT	DISULFID	163
FT	DISULFID	168
FT	DISULFID	185
FT	DISULFID	200
FT	DISULFID	230
FT	DISULFID	262
FT	DISULFID	292
FT	DISULFID	324
FT	DISULFID	354
FT	DISULFID	386
FT	DISULFID	416
FT	DISULFID	448
FT	DISULFID	478
FT	DISULFID	510
FT	DISULFID	540
FT	DISULFID	580
FT	DISULFID	610
FT	DISULFID	642
FT	DISULFID	672
FT	CARBOHYD	398
FT	CARBOHYD	603
FT	CARBOHYD	654
FT	CARBOHYD	661
FT	CARBOHYD	679
FT	LIPID	745
FT	SITE	756
FT	CONFLICT	724
FT	SEQUENCE	768 AA; 83098 MW; 65173074D2F66E8 CRC64; A -> E (IN REF. 2).

```

Query Match      68.3%; Score 455; DB 1; Length 768;
Best Local Similarity 67.5%; Pred. No. 3.4e-39;
Matches 79; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

Qy    1 WTYHSEKRMNOMRARRPCRDNYTDLVALONKAEELEYEKTLPFSRSYWGIRKGIGIW 60
Db    42 WTYNSTKAVSWMNSRYFCRRHFPTDVLVALONKNELAHLNDVYPFNSTSYWIGIRKNKW 101
Qy    61 TWVGNTSKSLTEBEAENMGDEGPNNKKNEEDCYEIYIKRNKDACKMWDDACHKLKALC 117
Db    102 TWWGNTKLTLEEAENWADNEPNKKNKNOCEIYIKSNAPSAGKMWDPECFKRRLAC 158

RESULT 11
LEM3_BOVIN
ID     LEM3_BOVIN          STANDARD:        PRT:       646 AA.
AC     P42201;
DT     01-NOV-1995 (Rel. 32, created)
DR     01-NOV-1995 (Rel. 32, last sequence update)
DT     15-JUL-1999 (Rel. 38, last annotation update)
DE     p-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEW)
DE     (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECMA3).
GN     SELF.
OS     Bos taurus (Bovine).
OC     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC     Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC     Bovidae; Bovinae; Bos.
RX     NCBI-TaxID=9913;
RN     [1]
RP     SEQUENCE FROM N.A.
RC     TISSUE-Capillary endothelium;
RX     MEDLINE=93249394; PubMed=7683458;
RA     Strabel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;
RT     "Isolation and characterization of a bovine cDNA encoding a
RT     functional homolog of human p-selectin."
RL     Biochem. Biophys. Res. Commun. 192:338-344(1993).
CC     -1 FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC     TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC     INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC     LEUKOCYTES. THE LIGAND RECOGNIZED IS STALKY-LEWIS X.
CC     CC
CC     -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC     CC
CC     -1 TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC     AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC     ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC     THE CELL SURFACE.
CC     CC
CC     -1 SIMILARITY: TO OTHER SELECTINS/LECTINS.
CC     CC
CC     -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC     CC
CC     -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC     CC
CC     -1 SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS. BOVINE P-LECTIN LACKS
CC     THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENDS.
-----
CC     This SWISS-PROT entry is copyright. It is produced through a collaboration
CC     between the Swiss Institute of Bioinformatics and the EMBL Outstation --
CC     the European Bioinformatics Institute. There are no restrictions on its
CC     use by non-profit institutions as long as its content is in no way
CC     modified and this statement is not removed. Usage by and for commercial
CC     entities requires a license agreement (see http://www.isdb.ch/announce/
CC     or send an email to license@isb-sib.ch).
-----
DB    EMBL, LI2041, AAA30743.1; -.
DR    HSHP, PI6109, IESB.
DR    InterPro, IPRO00561; EGF-like.
DR    InterPro, IPRO02396; Selectin.
DR    InterPro, IPRO00346; Sushi_SCR_CCP.
DR    InterPro, IPRO01304; lectin_c.
DR    Pfam, PF00008; EGF_1.
DR    Pfam, PF00059; Lectin_C_1.
DR    Pfam, PF00084; sushi_6.
DR    PRINTS, PR00343; SELECTIN.
DR    SMART, SM00032; CCP_6.
DR    SMART, SM00034; CLECT_1.
```

DR SMART: SM00181; EGF\_1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE; PS50041; C-TYPE LECTIN\_2; 1.  
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 Selectin; Signal; Sushi; Repeat.  
 FT SIGNAL 1 41  
 FT CHAIN 1 41  
 FT DOMAIN 42 646  
 FT TRANSMEM 588 611  
 FT DOMAIN 612 646  
 FT DOMAIN 58 158  
 FT DOMAIN 159 195  
 FT DOMAIN 199 258  
 FT DOMAIN 261 320  
 FT DOMAIN 323 382  
 FT DOMAIN 385 444  
 FT DOMAIN 457 516  
 FT DOMAIN 519 578  
 FT DOMAIN 58 158  
 FT DISULFID 131 150  
 FT DISULFID 163 174  
 FT DISULFID 168 183  
 FT DISULFID 185 194  
 FT DISULFID 200 244  
 FT DISULFID 230 257  
 FT DISULFID 262 306  
 FT DISULFID 292 319  
 FT DISULFID 324 368  
 FT DISULFID 354 381  
 FT DISULFID 386 430  
 FT DISULFID 416 443  
 FT DISULFID 458 502  
 FT DISULFID 488 515  
 FT DISULFID 520 564  
 FT DISULFID 550 577  
 FT CARBOHYD 48 48  
 FT CARBOHYD 54 54  
 FT CARBOHYD 80 80  
 FT CARBOHYD 180 180  
 FT CARBOHYD 212 212  
 FT CARBOHYD 219 219  
 FT CARBOHYD 336 336  
 FT CARBOHYD 481 481  
 FT CARBOHYD 532 532  
 FT CARBOHYD 539 539  
 FT CARBOHYD 557 557  
 FT SITE 634 637  
 SQ SEQUENCE 646 AA; 71229 MW; 573912A4627A6ACA CRC64;  
 Query Match 67.7%; Score 451; DB 1; Length 646;  
 Best local similarity 66.7%; Pred. No. 7.3e-39;  
 Matches 78; Conservative 10; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTHYSEKPMNORARFCEDNTDVAIONKAEIYELEKTFPSSSYWIGIRKIGIM 60  
 DB 42 WTHYHNKRYTSMVSRACFYCKYTTDLVAIONKNEIAYLNETTIPYNSYWIGIRKINNM 101  
 OY 61 TWVGTKKTLTEAEENMGDGEPPNKKNEKCEVETIKRNNDAGWMDACHKLKAALC 117  
 DB 102 TWVGTKKTLTEAEENWADNPPNKKRNNDCEVELYIKSLSPGKMWDEPCWKRRRALC 158  
 RESULT 12  
 LEM2\_BOVIN  
 ID LEM2\_BOVIN STANDARD; PRT; 485 AA.  
 AC P98107;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)

DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
 DE (CD62e).  
 GN SELE.  
 OS Bos taurus (Bovine).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 CC Bovidae; Bovinae; Bos.  
 CC NCBI\_Taxid=9913;  
 RN [1]  
 RP SQUENCE FROM N.A.  
 RC TISSUE-Adrenal gland;  
 RX MEDLINE=93382537; PubMed=7690465;  
 RA Nguyen M., Strubel N.A., Bischoff J.;  
 RT "A role for sialyl Lewis x/A glycoconjugates in capillary  
 morphogenesis";  
 RL Nature 365:267-269(1993).  
 CC -!- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
 CC POLYGLYCOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
 CC GLYCOLIPIDS).  
 CC -!- FUNCTION: INVOLVED IN CAPILLARY MORPHOGENESIS BY BINDING TO A  
 CC BOVINE CAPILLARY ENDOTHELIAL (BCE) CELL SIALYL LEWIS-X AND/OR  
 CC SIALYL LEWIS A-CONTAINING LIGAND (PROBABLY).  
 CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -!- SIMILARITY: TO OTHER SELECTINS/LECTINS.  
 CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -!- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAINS; BOVINE E-LECTIN LACKS  
 CC THE HUMAN SUSHI-4 AND 5 EQUIVALENTS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L12039; AAA02991.1; -  
 CC HSP: P16581; IESL.  
 DR DR HSP: P16581; IESL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; Lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sushi; 4.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 4.  
 DR SMART: SM00034; CLECT\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE LECTIN\_2; 1.  
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 Selectin; Signal; Sushi; Repeat.  
 FT SIGNAL 1 22  
 FT CHAIN 23 485  
 FT DOMAIN 23 430  
 FT TRANSMEM 431 453  
 FT DOMAIN 454 485  
 FT DOMAIN 39 139  
 FT DOMAIN 140 176  
 FT DOMAIN 180 238  
 FT DOMAIN 241 300  
 FT DOMAIN 303 363  
 FT DOMAIN 366 422  
 FT DISULFID 41 139  
 FT DISULFID 112 131  
 FT DISULFID 144 155

```
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 166 175 BY SIMILARITY.
FT DISULFID 181 224 BY SIMILARITY.
FT DISULFID 210 237 BY SIMILARITY.
FT DISULFID 242 286 BY SIMILARITY.
FT DISULFID 272 299 BY SIMILARITY.
FT DISULFID 304 349 BY SIMILARITY.
FT DISULFID 335 362 BY SIMILARITY.
FT DISULFID 367 408 BY SIMILARITY.
FT DISULFID 394 421 BY SIMILARITY.
FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 485 AA; 53200 MW; AE931C9B521E3904 CRC64;

Query Match 66.8%; Score 445; DB 1; Length 485;
Best Local Similarity 65.8%; Pred. No. 2,2e-38;
Matches 77; Conservative 14; Mismatches 26; Indels 0; Gaps 0;

Oy 1 WTHYSEPMNQRRRCRDNYTDLVAIONKAELEYLEKTLPSRSYWGIRKIGTW 60
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 23 WSYHSTEMTEFEERADYCKQTYTALVAIONDEELEYLYNTEFSYSPSYWIGIRKINGTW 82
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

Oy 61 TWVGNSKSLTEEAENMGDEPNKKKEDCVETIKRNKDGAKNDMDCHKIKALC 117
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 83 TWIGNSKSLTEEAENMGDEPNKKKEDCVETIKRNKDGAKNDMDCHKIKALC 139
    1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 13
LEM3_HUMAN STANDARD: PRT; 830 AA.
ID LEM3_HUMAN STANDARD: PRT; 830 AA.
AC P16109.
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN SELP OR GMRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=89168432; PubMed=2466574;
RA Johnston G.I., Cook R.G., McEvers R.P.;
RT "Cloning of GMP-140, a granule membrane protein of platelets and
RT endothelium: sequence similarity to proteins involved in cell
RT adhesion and inflammation.";
RL Cell 56:1033-1044(1989).
[2]
RN PALMITOYLATION.
RP MEDLINE=93265599; PubMed=7684381;
RA Fujimoto T., Stroud E., Whalley R.E., Prescott S.M., Muszbek L.,
RA Laposata M., McEvers R.P.;
RT "P-selectin is acylated with palmitic acid and stearic acid at
RT cysteine 766 through a thioester linkage.";
RL J. Biol. Chem. 268:11394-11400(1993).
[3]
RN STRUCTURE BY NMR OF 160-199.
RP MEDLINE=97057176; PubMed=8901515;
RA Freedman S.J., Sanford D.G., Bachovichin W.W., Furie B.C., Baley J.D.,
RA Furie B.;
RT "Structure and function of the epidermal growth factor domain of P-
RT selectin.";
```

```
RL Biochemistry 35:13733-13744(1996).
[4]
RN 3D-STRUCTURE MODELING OF 42-161.
RP MEDLINE=94093388; PubMed=7505680;
RX Bajorath J., Stenkamp R., Arutfo A.;
RT "Knowledge-based model building of proteins: concepts and examples.";
RL Protein Sci. 2:1798-1810(1993).
[5]
RN VARIANTS ASN-331; ASP-603; VAL-640 AND PRO-756.
RX MEDLINE=98334547; PubMed=9668170;
RA Herrmann S.M., Ricard S., Nicaud V., Mallet C., Evans A.,
RA Ruidavets J.B., Arveiler D., Luc G., Cambien F.;
RT "The P-selectin gene is highly polymorphic: reduced frequency of the
RT Pro715 allele carriers in patients with myocardial infarction.";
RL Hum. Mol. Genet. 7:1277-1284(1998).
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC THE CELL SURFACE.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 9 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME=PRO; NOTE=CD guide CD62P entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd62p.htm".
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-stb.ch/announce/
or send an email to license@isb-stb.ch).
-----
CC EMBL; M60234; AAA35910.1; -.
CC EMBL; M60217; AAA35910.1; JOINED.
CC EMBL; M60218; AAA35910.1; JOINED.
CC EMBL; M60219; AAA35910.1; JOINED.
CC EMBL; M60222; AAA35910.1; JOINED.
CC EMBL; M60223; AAA35910.1; JOINED.
CC EMBL; M60224; AAA35910.1; JOINED.
CC EMBL; M60225; AAA35910.1; JOINED.
CC EMBL; M60226; AAA35910.1; JOINED.
CC EMBL; M60227; AAA35910.1; JOINED.
CC EMBL; M60228; AAA35910.1; JOINED.
CC EMBL; M60229; AAA35910.1; JOINED.
CC EMBL; M60231; AAA35910.1; JOINED.
CC EMBL; M60232; AAA35910.1; JOINED.
CC EMBL; M60233; AAA35910.1; JOINED.
CC EMBL; M25322; AAA35911.1; -.
CC PIR; A30359; AAA35911.1; -.
CC PDB; 1FSB; 01-APR-97.
CC PDB; 1KUD; 03-APR-96.
CC MIM; 173610; -.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushl_SCR_CCP.
DR InterPro: IPR001304; Ilectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00084; Ilectin_C_1.
DR Pfam: PF00059; Ilectin_C_1.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP_9.
DR SMART; SK00034; CLECT; 1.
DR SMART; SK00181; EGF_1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C_TYPE_LLECTIN_1; 1.
```





```

FT DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 710 733 POTENTIAL.
FT DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 258 SUSH1.
FT DOMAIN 261 320 SUSH1 1.
FT DOMAIN 323 382 SUSH1 2.
FT DOMAIN 385 444 SUSH1 3.
FT DOMAIN 447 506 SUSH1 4.
FT DOMAIN 509 568 SUSH1 5.
FT DOMAIN 579 638 SUSH1 6.
FT DOMAIN 641 700 SUSH1 7.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 324 319 BY SIMILARITY.
FT DISULFID 354 368 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 448 492 BY SIMILARITY.
FT DISULFID 478 505 BY SIMILARITY.
FT DISULFID 510 554 BY SIMILARITY.
FT DISULFID 540 567 BY SIMILARITY.
FT DISULFID 580 624 BY SIMILARITY.
FT DISULFID 610 637 BY SIMILARITY.
FT DISULFID 642 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT CARBOHYD 45 45 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 654 654 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 661 661 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAG. . .) (POTENTIAL).
FT LIPID 745 745 PALMITATE (BY SIMILARITY).
FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
SQ SEQUENCE 768 AA; 83517 MW; 266D7E8A5F3F1316 CRC64;

Query Match 66.4%; Score 442; DB 1; Length 768;
Best Local Similarity 64.1%; Pred. No. 7.5e-38;
Matches 75; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQARARPCRDNYTDVAIONKAEIYELEKTLFSPRSYVWIGIRKIGIW 60
DB 42 WTYNSTAYSNNSRACRKHFTDLVAIONKNEIHLNDVAPYNSYVWIGIRKINWK 101
QY 61 TWGYNKSLTEBAENWGGCEPNNKKNKEDCEVETIKRKNDKAKWDACHTKALALC 117
DB 102 TWGYNKSLTEBAENWGGCEPNNKKNKEDCEVETIKRKNDKAKWDACHTKALALC 158

RESULT 15
LEM2_RABIT STANDARD: PRT; 551 AA.
AC P27113;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E-selectin precursor (endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE.
OS Oryctolagus cuniculus (Rabbit).
```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_Taxid=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=9218729; PubMed=1372169;
RA Larijan J.D., Tsang T.C., Rumberger J.M., Burns D.K.;
RT "Characterization of cDNA and genomic sequences encoding rabbit
RT ELAM-1: conservation of structure and functional interactions with
RT leukocytes."
RL DNA Cell Biol. 11:149-162(1992).
CC -! FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS SIMILY-LEWIS X (ALPHA1-3)FUCCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -! SUBCELLULAR LOCATION: Type I membrane protein.
CC -! INDUCTION: BY CYTOKINES.
CC -! SIMILARITY: TO OTHER SELECTINS/LECTINS.
CC -! SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -! SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -! SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M91004; AAA31243.1; -
DR EMBL: M91005; AAA31244.1; -
DR HSSP: P16581; IKJA.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sush1; 5.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 5.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_2; 1.
DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT CHAIN 1 23
FT DOMAIN 24 551 E-SELECTIN.
FT DOMAIN 24 495 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 496 517 POTENTIAL.
FT DOMAIN 518 551 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 40 140 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 141 177 EGF-LIKE.
FT DOMAIN 181 240 SUSH1 1.
FT DOMAIN 243 302 SUSH1 2.
FT DOMAIN 305 365 SUSH1 3.
FT DOMAIN 368 428 SUSH1 4.
FT DOMAIN 431 487 SUSH1 5.
FT DISULFID 42 140 BY SIMILARITY.
FT DISULFID 113 132 BY SIMILARITY.
FT DISULFID 145 156 BY SIMILARITY.
FT DISULFID 150 165 BY SIMILARITY.
FT DISULFID 167 176 BY SIMILARITY.
FT DISULFID 182 226 BY SIMILARITY.
FT DISULFID 212 239 BY SIMILARITY.
FT DISULFID 244 288 BY SIMILARITY.
FT DISULFID 274 301 BY SIMILARITY.
```



```

FT CONFLICT 327 327 K -> N (IN REF. 2).
FT CONFLICT 363 363 V -> A (IN REF. 2).
FT CONFLICT 384 384 V -> M (IN REF. 2).
FT CONFLICT 461 484 KFPSSESSECLQPGSGTQMPEDLI ->
FT SEQUENCE 484 AA; 52567 MM; AEF74FE25C1FD013 CRC64;
SQ

Query Match 63.88; Score 425; DB 1; Length 484;
Best Local Similarity 63.28; Pred. No. 2.5e-36;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY 1 WYHSEKPMQMRARRCRDNYTDLVAIONKAELEYLEKTLFGRSYWIGIRKIGTW 60
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 23 WSYSTSTETMTFDASATCCOQRTHLVAIONHAELEYLSTFNYSASTYIGIRKINGTW 82
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
QY 61 TWVGTNKSILTEAEENWGDGEPNNKKNKEDCEIYIKRNKADGKWDNDCKHLKALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 83 TWIGTKKALTEATWMAWGPENNNKQSDCEIYIKRDKSGKWNDECSKKKLLALC 139
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 17
LEM2_HUMAN
ID LEM2_HUMAN STANDARD; PRT; 610 AA.
AC P16581; P16111;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE E-selectin precursor (endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
GN SELE OR ELAM1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=90175359; PubMed=1689848;
RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C.,
RA Parek M., Pltack C., Tizard R., Goetz S., McCarthy K., Hopple S.,
RA Lobb R.;
RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning
RT and functional interactions.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).
RN 12
RP SEQUENCE FROM N.A.
RX MEDLINE=89162047; PubMed=2466335;
RA Bevilacqua M.P., Stengelin S., Gmbrone M.A. Jr., Seed B.;
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for
RT neutrophils related to complement regulatory proteins and lectins.";
RL Science 243:1160-1165(1989).
RN 13
RP SEQUENCE FROM N.A.
RX MEDLINE=9115870; PubMed=1703529;
RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,
RA Gmbrone M.A. Jr., Bevilacqua M.P.;
RT "Structure and chromosomal location of the gene for endothelial-
RT leukocyte adhesion molecule 1.";
RL J. Biol. Chem. 266:2466-2473(1991).
RN 14
RP LIGAND.
RX MEDLINE=91068005; PubMed=1701274;
RA Phillips M.L., Nudelma E., Gaeta F.C., Perez M., Singhal A.K.,
RA Hakomori S., Paulson J.C.;
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate
RT ligand, sialyl-Lex.";
RL Science 250:1130-1132(1990).
RN 15
RP 3D-STRUCTURE MODELLING OF LECTIN DOMAIN.
RX MEDLINE=93202275; PubMed=7681016;
RA Mills A.;
RT "Modelling the carbohydrate recognition domain of human E-selectin.";
```

```

RL FEBS Lett. 319:5-11(1993).
RN [6]
RN X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.
RX MEDLINE=94150646; PubMed=7509040;
RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,
RA Huang K.-S., Presky D.H., Familletti P.C., Wollitzky B.A., Burns D.K.;
RT "Insight into E-selectin/ligand interaction from the crystal
RT structure and mutagenesis of the Lec/EGF domains.";
RL Nature 367:532-538(1994).
RN [7]
RN VARIANT ARG-149.
RX MEDLINE=95179107; PubMed=7533025;
RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,
RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;
RT "E-selectin polymorphism and atherosclerosis: an association study.";
RL Hum. Mol. Genet. 3:1935-1937(1994).
RN [8]
RN VARIANT ARG-149.
RX MEDLINE=99134508; PubMed=9933738;
RA Ye S.Q., Usher D., Virgil D., Zhang L.Q., Yochim S.E., Gupta R.;
RT "A PstI polymorphism detects the mutation of serine-128 to arginine in
RT CD 62E gene - a risk factor for coronary artery disease.";
RL J. Biomed. Sci. 6:18-21(1999).
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A
CC HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY
CC HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH
CC ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN
CC UNSELECTED POPULATION (SER-149).
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD62E entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62e.htm".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30640; AAA52377.1; -;
DR EMBL; M61893; AAA52375.1; JOINED.
DR EMBL; M61895; AAA52375.1; JOINED.
DR EMBL; M61887; AAA52375.1; JOINED.
DR EMBL; M61888; AAA52375.1; JOINED.
DR EMBL; M61890; AAA52375.1; JOINED.
DR EMBL; M61891; AAA52375.1; JOINED.
DR EMBL; M61892; AAA52375.1; JOINED.
DR EMBL; M24736; AAA52376.1; -;
DR PIR; A32606; A32606.
DR PIR; A35046; A35046.
DR PIR; A38615; A38615.
DR PDB; 1EST; 31-AUG-94.
DR PDB; 1KJA; 03-APR-96.
DR MIM; 131210; -;
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sushl_SCR_CCP.
DR InterPro; IPR001304; Lectin_C.
DR Pfam; PF00008; EGF_1.
DR Pfam; PF00059; Lectin_C_1.
DR Pfam; PF00084; sushl_6.
DR PRINTS; PR00343; SELECTIN.
DR SMART; SM00032; CCP; 6.
```

```

DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat; Polymorphism; 3d-structure.
FT SIGNAL 1 21
FT CHAIN 22 610
FT DOMAIN 22 556
FT TRANSMEM 557 578
FT DOMAIN 579 610
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 238
FT DOMAIN 241 300
FT DOMAIN 303 363
FT DOMAIN 366 426
FT DOMAIN 429 489
FT DOMAIN 492 548
FT DISULFID 40 138
FT DISULFID 111 130
FT DISULFID 143 154
FT DISULFID 148 163
FT DISULFID 165 174
FT DISULFID 180 224
FT DISULFID 210 237
FT DISULFID 242 286
FT DISULFID 272 299
FT DISULFID 304 349
FT DISULFID 335 362
FT DISULFID 367 412
FT DISULFID 398 425
FT DISULFID 430 475
FT DISULFID 461 488
FT DISULFID 493 534
FT DISULFID 520 547
FT CARBOHYD 25 25
FT CARBOHYD 145 145
FT CARBOHYD 160 160
FT CARBOHYD 179 179
FT CARBOHYD 199 199
FT CARBOHYD 203 203
FT CARBOHYD 265 265
FT CARBOHYD 312 312
FT CARBOHYD 332 332
FT CARBOHYD 503 503
FT CARBOHYD 527 527
FT VARIANT 130 130
FT VARIANT 149 149
FT VARIANT 149 149
FT VARIANT 295 295
FT VARIANT 421 421
FT VARIANT 468 468
FT VARIANT 575 575
FT SEQUENCE 610 AA; 66655 MW; 7043E3C0D1229229 CRC64;

Query Match 62.3%; Score 415; DB 1; Length 610;
Best Local Similarity 60.7%; Pred. No. 3-5e-35;
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

```

```

QY 1 WYHYHSEKPMNORARPCRDNYTDVAIQNAEIEYLEKTLPEFSRYWIGIRKIGIM 60
DB 22 WSYNTSTAMFYDEASAYCOQRTHVAIQNKEIEYLSILSYSSYWIGIRKNNW 81

```

```

QY 61 TWGTSNKSLTEAEENMGDEPNKKNKEDCEYIYIKRKNDAKWNDACHKLKAALC 117
DB 82 VWVGTQRPILTEAKNMAKAPGNPNKQKDECEYIYIKRKNDAKWNDACHKLKAALC 138

RESULT 18
LEW2_RAT
ID LEW2_RAT STANDARD; PRT; 549 AA.
AC P98105;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E)
CN SELE OR ELAM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUCOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS STAYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: I25527; AAA4113.1; -.
CC DR HSSP: P16581; IKTA.
CC DR InterPro: IPR000561; EGF-like.
CC DR InterPro: IPR000436; Sushi_SCR_CCP.
CC DR InterPro: IPR001304; Lectin_c.
CC DR Pfam: PF00008; EGF_1.
CC DR Pfam: PF00059; lectin_c; 1.
CC DR Pfam: PF00084; sushi; 5.
CC DR SMART: SM00032; CCP; 5.
CC DR SMART: SM00034; CLECT; 1.
CC DR SMART: SM00181; EGF; 1.
CC DR PROSITE: PS00022; EGF_1; 1.
CC DR PROSITE: PS01186; EGF_2; 1.
CC DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
CC DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
CC KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 549
FT DOMAIN 22 494
FT TRANSMEM 495 516
FT DOMAIN 517 549
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 486
FT DOMAIN 430 486
FT SUSHI 5.

```



```

DB 23 WYNNASTEAMTDEASTYQOQRTHLVAIONOEIKYLSMFTYPTTYWIGIRVNNKW 82
OY 61 TWVGTNKSILTEAEANWGDPEPNKKNECCEVEIYIKRNKDACWMDACHKIKALC 117
DB 83 TWIGTOKLITEAKWNAPEBPNNKKNDECCVEIYIKRDKSGKMDERCKRKKALC 139

RESULT 20
ID LEM2_MOUSE STANDARD; PRT; 612 AA.
AC 000690:
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DE 01-FEB-1996 (Rel. 33, Last annotation update)
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62E).
CN SELE OR ELAM-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RA MEDLINE=92283265; PubMed=1375914;
RA Becker-Andre M., van Huijsduijn R.H., Losberger C., Whelan J.,
RA Delamarier J.F.;
RT "Murine endothelial leukocyte-adhesion molecule 1 is a close
RT structural and functional homologue of the human protein.";
RL Eur. J. Biochem. 206:401-411(1992).
RN 12
RP SEQUENCE FROM N.A.
RA MEDLINE=92340571; PubMed=1378846;
RA Weller A., Isenmann S., Vestweber D.;
RT "Cloning of the mouse endothelial selectins. Expression of both E-
RT and P-selectin is inducible by tumor necrosis factor alpha.";
RL J. Biol. Chem. 267:15176-15183(1992).
CC 1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS STAYL-L-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF
CC POLYACETOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOLIPIDS).
CC 1- SUBCELLULAR LOCATION: Type I membrane protein.
CC 1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC 1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC 1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC 1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80778; AAA37547.1; .
DR EMBL: M87862; AAA37577.1; ALT_INIT.
DR HSSP: P16581; IKAJ.
DR MGD: MGI:98278; Sele.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_c_1.
DR Pfam: PF00084; sushi_6.
DR PRINTS: PR000343; SELECTIN.
DR SMART: SM00032; CCP; 6.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.

```

```

DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS00041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 21
FT CHAIN 22 612
FT DOMAIN 22 557
FT TRANSMEM 558 579
FT DOMAIN 580 612
FT DOMAIN 38 138
FT DOMAIN 139 175
FT DOMAIN 179 239
FT DOMAIN 242 301
FT DOMAIN 304 364
FT DOMAIN 367 427
FT DOMAIN 430 490
FT DOMAIN 493 549
FT DISULFID 40 138
FT DISULFID 111 130
FT DISULFID 143 154
FT DISULFID 148 163
FT DISULFID 165 174
FT DISULFID 180 225
FT DISULFID 210 238
FT DISULFID 243 287
FT DISULFID 273 300
FT DISULFID 305 350
FT DISULFID 336 363
FT DISULFID 368 413
FT DISULFID 399 426
FT DISULFID 431 476
FT DISULFID 462 489
FT DISULFID 494 535
FT DISULFID 521 548
FT CARBOHYD 25 25
FT CARBOHYD 145 145
FT CARBOHYD 192 192
FT CARBOHYD 203 203
FT CARBOHYD 266 266
FT CARBOHYD 333 333
FT CARBOHYD 391 391
FT CARBOHYD 504 504
FT CARBOHYD 528 528
SQ SEQUENCE 612 AA; 66749 MW; 86F05713F0EC3C3D CRC64;

Query Match 60.8%; Score 405; DB 1; Length 612;
Best Local Similarity 59.8%; Pred. No. 3, 7e-34;
Matches 70; Conservative 14; Mismatches 33; Indels 0; Gaps 0;

OY 1 WYHYSEKPPNMQRRARFCRDNYTDIVAIQNAKEIYLEKTELPFSRSYYWIGIRKIGTW 60
DB 22 WYNNASSELMTIDEASAYCQRDTHLVAIONKEEYINLSNKHSPSYWIGIRVNNW 81
OY 61 TWVGTNKSILTEAEANWGDPEPNKKNECCEVEIYIKRNKDACWMDACHKIKALC 117
DB 82 TWIGTOKLITEAKWNAPEBPNNKKNDECCVEIYIKRDKSGKMDERCKRKKALC 139

RESULT 21
ID MANR_HUMAN STANDARD; PRT; 1456 AA.
AC P22897;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Macrophage mannose receptor precursor (MMR) (CD206 antigen).
GN MRC1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;

```

```

RN [1] SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RP TISSUE=Placenta;
RA MEDLINE=90324192; PubMed=2373685;
RT Taylor M.E., Conary J.T., Lemnarz W.R., Stahl P.D., Dickamer K.;
RT "Primary structure of the mannose receptor contains multiple motifs
RT resembling carbohydrate-recognition domains.";
RN J. Biol. Chem. 265:12156-12162(1990).
RN (2)
RP SEQUENCE FROM N.A.
RA MEDLINE=93052405; PubMed=1294118;
RT Kim S.J., Ruiz N., Bezouska K., Dickamer K.;
RT "Organization of the gene encoding the human macrophage mannose
RT receptor (MRC1).";
RN Genomics 14:721-727(1992).
RN [3]
RP STUDIES ON THE BINDING OF INDIVIDUAL LECTIN DOMAINS.
RX MEDLINE=92112893; PubMed=1730714;
RA Taylor M.E., Bezouska K., Dickamer K.;
RT "Contribution to ligand binding by multiple carbohydrate-recognition
RT domains in the macrophage mannose receptor.";
RN J. Biol. Chem. 267:1719-1726(1992).
RN [4]
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 642-788.
RA MEDLINE=20347275; PubMed=10779515;
RA Feinberg H., Park-Snyder S., Kolatkar A.R., Heise C.T., Taylor M.E.,
RA Weis W.I.;
RT "Structure of a C-type carbohydrate recognition domain from the
RT macrophage mannose receptor.";
RN J. Biol. Chem. 275:21539-21548(2000).
RN [5]
RT -1 FUNCTION: MEDIATES THE ENDOCYTOSIS OF GLYCOPROTEINS BY
RN MACROPHAGES, IN SEVERAL RECOGNITION AND UPTAKE PROCESSES.
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.
CC -1 MISCELLANEOUS: CRDS 1-3 HAVE AT MOST VERY WEAK AFFINITY FOR
CC CARBOHYDRATE. CRD 4 SHOWS THE HIGHEST AFFINITY BINDING AND HAS
CC MULTISPECIFICITY FOR A VARIETY OF MONOSACCHARIDES. AT LEAST 3 CRDS
CC (4, 5, AND 7) ARE REQUIRED FOR HIGH AFFINITY BINDING AND
CC ENDOCYTOSIS OF MULTIVALENT GLYCOCONGUGATES.
CC -1 SIMILARITY: CONTAINS 8 C-TYPE LECTIN FAMILY DOMAINS.
CC -1 SIMILARITY: CONTAINS 1 RICIN B-TYPE LECTIN DOMAIN.
CC -1 DATABASE: NAME=PROX; NOTE=CD guide CD206 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/prox/guide/1044341355-5.htm".
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; J05550; AAA59868.1; -;
DR EMBL; M93221; AAA60389.1; -;
DR EMBL; M93192; AAA60389.1; JOINED.
DR EMBL; M93193; AAA60389.1; JOINED.
DR EMBL; M93194; AAA60389.1; JOINED.
DR EMBL; M93195; AAA60389.1; JOINED.
DR EMBL; M93196; AAA60389.1; JOINED.
DR EMBL; M93197; AAA60389.1; JOINED.
DR EMBL; M93198; AAA60389.1; JOINED.
DR EMBL; M93199; AAA60389.1; JOINED.
DR EMBL; M93200; AAA60389.1; JOINED.
DR EMBL; M93201; AAA60389.1; JOINED.
DR EMBL; M93202; AAA60389.1; JOINED.
DR EMBL; M93203; AAA60389.1; JOINED.
DR EMBL; M93204; AAA60389.1; JOINED.
DR EMBL; M93205; AAA60389.1; JOINED.
DR EMBL; M93206; AAA60389.1; JOINED.
DR EMBL; M93207; AAA60389.1; JOINED.
DR EMBL; M93208; AAA60389.1; JOINED.
DR EMBL; M93209; AAA60389.1; JOINED.
DR EMBL; M93210; AAA60389.1; JOINED.
DR EMBL; M93211; AAA60389.1; JOINED.

```

DR	EMBL:	M93212;	AAA60389.1;	JOINED.	
DR	EMBL:	M93213;	AAA60389.1;	JOINED.	
DR	EMBL:	M93214;	AAA60389.1;	JOINED.	
DR	EMBL:	M93215;	AAA60389.1;	JOINED.	
DR	EMBL:	M93216;	AAA60389.1;	JOINED.	
DR	EMBL:	M93217;	AAA60389.1;	JOINED.	
DR	EMBL:	M93218;	AAA60389.1;	JOINED.	
DR	EMBL:	M93219;	AAA60389.1;	JOINED.	
DR	EMBL:	M93220;	AAA60389.1;	JOINED.	
DR	PIR:	A36563;	A36563.		
DR	PDB:	1EGG;	30-AUG-00.		
DR	PDB:	1EGI;	30-AUG-00.		
DR	MIM:	153618;	-		
DR	InterPro:	IPR000562;	FN_Type_II.		
DR	InterPro:	IPR000772;	Ricin_B_lectin.		
DR	InterPro:	IPR001304;	lectin_c.		
DR	Pfam:	PF000400;	fn2; 1.		
DR	Pfam:	PF000059;	lectin_c; 8.		
DR	Pfam:	PF00652;	Ricin_B_lectin; 1.		
DR	PRINTS:	PR00013;	FNTypePII.		
DR	ProDom:	PD000995;	FN_Type_II; 1.		
DR	SMART:	SM00034;	CLECT; 8.		
DR	SMART:	SM00059;	FN2; 1.		
DR	SMART:	SM00458;	RICIN; 1.		
DR	PROSITE:	PS00615;	C_Type_LECTIN_1; 6.		
DR	PROSITE:	PS50041;	C_Type_LECTIN_2; 8.		
DR	PROSITE:	PS00023;	FIBRONECTIN_2; 1.		
DR	PROSITE:	PS50231;	RICIN_B_lectin; 1.		
KW	Receptor:	Signal;	Calcium-binding; Transmembrane; Repeat;		
KW	Glycoprotein:	Endocytosis; Macrophage; Lectin; Antigen; 3D-structure.			
FT	SIGNAL	1	18	POTENTIAL.	
FT	CHAIN	19	1456	MACROPHAGE MANNOSE RECEPTOR.	
FT	DOMAIN	19	1383	EXTRACELLULAR (POTENTIAL).	
FT	TRANSMEM	1384	1411	POTENTIAL.	
FT	DOMAIN	1412	1456	CYTOPLASMIC (POTENTIAL).	
FT	DOMAIN	22	142	RICIN B-TYPE LECTIN.	
FT	DOMAIN	157	212	FIBRONECTIN TYPE-II.	
FT	DOMAIN	216	344	C-TYPE LECTIN 1 (LONG FORM).	
FT	DOMAIN	360	490	C-TYPE LECTIN 2 (LONG FORM).	
FT	DOMAIN	502	629	C-TYPE LECTIN 3 (LONG FORM).	
FT	DOMAIN	644	781	C-TYPE LECTIN 4 (LONG FORM).	
FT	DOMAIN	805	926	C-TYPE LECTIN 5 (LONG FORM).	
FT	DOMAIN	943	1083	C-TYPE LECTIN 6 (LONG FORM).	
FT	DOMAIN	1100	1216	C-TYPE LECTIN 7 (LONG FORM).	
FT	DOMAIN	1228	1359	C-TYPE LECTIN 8 (LONG FORM).	
FT	DISULFID	646	659		
FT	DISULFID	680	777		
FT	DISULFID	753	769		
FT	CARBOHYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	344	344	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	529	529	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	926	926	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	930	930	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1160	1160	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1205	1205	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1265	1265	N-LINKED (GLCNAC. . .) (POTENTIAL).	
FT	CARBOHYD	1311	1311	N-LINKED (GLCNAC. . .) (POTENTIAL).	
SO	SEQUENCE	1456 AA;	166011 MW;	2645A5E3C576A5E3 CRC64;	

Query Match	23.6%	Score 157	DB 1	Length 1456	
Best Local Similarity	30.8%	Pred. No. 3,4e-08			
Matches	37	Conservative	29	Mismatches 44	Indels 10
				Gaps	4
QY	1	WTYHYSEKPMNQARRFCRDNYTDLVAIONAEIYELEKTL--PFSRSYWGIR-KIG	57		
DB	810	YQYVFSKEKETMDNARAFCKRNGPDVLSIQSEKKFLMKYVNRNDAGSAVFIGLLISD	869		
QY	58	GIWTYVGTKRSLTEKENNGDGEPPNNKKKKKEOCVELYIRKKNDDAGKWNDDACHKLKAALC	117		
DB	870	KKFAWMDGK---VDYVSWATGEPPNANEDCEVTVY---SNSGFWNDINCGYPNATFC	922		
RESULT	22				





RESULT 23  
MMGL\_MOUSE STANDARD: PRT; 304 AA.

ID MMGL\_MOUSE  
AC P49300;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-BP) (Macrophage galactose/N-acetylgalactosamine-specific lectin) (MMGL).  
GN MGL.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C3H/HEJ;  
RX MEDLINE=92268032; PubMed=1587794;  
RA Sato M., Kawakami K., Osawa T., Toyoshima S.;  
RT "Molecular cloning and expression of cDNA encoding a galactose/N-acetylgalactosamine-specific lectin on mouse tumoricidal macrophages.";  
RT J. Biochem. 111:331-336(1992).  
RN [2]  
RP SEQUENCE OF 102-120 AND 137-151.  
RC STRAIN=C3H/HEJ;  
RX MEDLINE=89197865; PubMed=3241002;  
RA Oda S., Sato M., Toyoshima S., Osawa T.;  
RT "Purification and characterization of a lectin-like molecule specific for galactose/N-acetyl-galactosamine from tumoricidal macrophages.";  
RT J. Biochem. 104:600-605(1988).  
CC -1- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. MAY PARTICIPATE IN THE INTERACTION BETWEEN TUMORICIDAL MACROPHAGES AND TUMOR CELLS.  
CC -1- SUBUNIT: HOMO-OLIGOMER.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- TISSUE SPECIFICITY: IS EXPRESSED ON THE SURFACE OF ACTIVATED MACROPHAGES.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: S36676; AAB22171.1; -;  
DR HSSP: P06734; 1KJE.  
DR MGD: MGI:96975; MGI.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT; 1.  
DR PROSITE: PS00615; C-TYPE-LECTIN\_1; 1.  
DR PROSITE: PS00441; C-TYPE-LECTIN\_2; 1.  
KW lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.  
FT DOMAIN 1 35 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 57 304 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 172 298 C-TYPE LECTIN (LONG FORM).  
FT DISULFID 173 184 BY SIMILARITY.  
FT DISULFID 201 296 BY SIMILARITY.  
FT DISULFID 274 288 BY SIMILARITY.  
FT CARBOHYD 74 74 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 166 166 N-LINKED (GLCNAc. . .) (POTENTIAL).  
SQ SEQUENCE 304 AA; 34596 MW; 3F79CD12C34F5BCC CRC64;

Query Match 23.0%; Score 153.5; DB 1; Length 304;  
Best Local Similarity 29.5%; Pred. NO. 1.3e-08;

Matches 38; Conservative 18; Mismatches 52; Indels 21; Gaps 6;

OY 1 WTTH-----YSEKPMNORARFCNDNTDLVAIONKAIELEKTLPESSYVIGIR 54  
DB 177 WTEHGGSCYWFSESEKSWPEADKYCRLNSHLVYNSLEONLQRL--ANVSWIGLT 234  
OY 55 KIGGIWTVW-GNFKSLTEEAENMGDGEPPN-----KKNEDEVEIYIKRNKQAGKNDA 108  
DB 235 DONGPWRWDGID--FEKFKNAPLQPDNWFHGIGGEDCAHI-----TTGGPWNDDY 287  
OY 109 CHKLKALC 117  
DB 288 CQTFEWMIC 296

RESULT 24  
KCCR\_MOUSE STANDARD: PRT; 548 AA.

ID KCCR\_MOUSE  
AC P70194;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Kupfer cell receptor.  
GN KCLR.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxId=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BA16/C; TISSUE=Liver;  
RA Takazawa R., Magatsuma H., Nomoto C., Watanabe Y., Akaike T.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCOSE.  
CC COULD BE INVOLVED IN ENDOCYTOSIS.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- TISSUE SPECIFICITY: KUPFER CELLS.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D88577; BAA13647.1; -;  
DR HSSP: P20693; 1HLJ.  
DR MGD: MGI:1859834; KCLR.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT; 1.  
DR PROSITE: PS00615; C-TYPE-LECTIN\_1; 1.  
DR PROSITE: PS00441; C-TYPE-LECTIN\_2; 1.  
KW Receptor; Transmembrane; Glycoprotein; Lectin; Signal-anchor.  
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT DOMAIN 70 548 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 438 538 C-TYPE LECTIN (SHORT FORM).  
FT DISULFID 440 536 BY SIMILARITY.  
FT DISULFID 516 528 BY SIMILARITY.  
FT CARBOHYD 86 86 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 92 92 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 115 115 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 132 132 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 209 209 N-LINKED (GLCNAc. . .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAc. . .) (POTENTIAL).  
SQ SEQUENCE 548 AA; 61268 MW; 6F6495EB20E73BD9 CRC64;





OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RX SEQUENCE FROM N.A.  
 RX MEDLINE=87257885; PubMed=3600647;  
 RA McPhaul M., Berg P.;  
 RT "Identification and characterization of cDNA clones encoding two  
 RT homologous proteins that are part of the asialoglycoprotein  
 RT receptor";  
 RL Mol. Cell. Biol. 7:1841-1847(1987).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=87250656; PubMed=3597443;  
 RA Halberg D.F., Wager R.E., Farrell D.C., Hildreth J., Quesenberry M.S.,  
 RA Loeb J.A., Holland E.C., Drickamer K.;  
 RT "Major and minor forms of the rat liver asialoglycoprotein receptor  
 RT are independent galactose-binding proteins. Primary structure and  
 RT glycosylation heterogeneity of minor receptor forms";  
 RL J. Biol. Chem. 262:9828-9838(1987).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RP STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;  
 RX MEDLINE=89170119; PubMed=3234178;  
 RA Sanford J.P., Elliott R.W., Doyle D.;  
 RT "Asialoglycoprotein receptor genes are linked on chromosome 11 in the  
 RT mouse";  
 RL DNA 7:721-728(1988).  
 RN [4]  
 RP SEQUENCE OF 201-301.  
 RX MEDLINE=8411554; PubMed=6319386;  
 RA Drickamer K., Mamou J.F., Bins G., Leung J.O.;  
 RT "Primary structure of the rat liver asialoglycoprotein receptor.  
 RT Structural evidence for multiple polypeptide species";  
 RL J. Biol. Chem. 259:770-778(1984).  
 CC -1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO  
 CC WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX  
 CC CARBOHYDRATE MOIETIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES  
 CC TERMINAL GALACTOSE AND N-ACETYLGLACTOSAMINE UNITS. AFTER LIGAND  
 CC BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND  
 CC TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE  
 CC DISSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE  
 CC SURFACE.  
 CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL  
 CC CELLS.  
 CC -1- MISCELLANEOUS: CALCIUM IS REQUIRED FOR LIGAND BINDING.  
 CC -1- MISCELLANEOUS: TWO TYPES OF RAT HEPATIC LECTIN HAVE BEEN  
 CC IDENTIFIED, RHL-1 AND RHL-2/3, HAVING A RELATIVE ABUNDANCE OF 4:1.  
 CC RHL-2 AND RHL-3 ONLY DIFFERS IN THEIR CARBOHYDRATE STRUCTURES.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 DR EMBL: M16347; AAA42038.1; -;  
 DR EMBL: J02762; AAA41522.1; -;  
 DR EMBL: X07636; CAA30476.1; -;  
 DR PIR: A25417; INRT2.  
 DR PIR: A26888; A28462.  
 DR PIR: A28462; A28462.  
 DR PIR: A31601; A31601.  
 DR HSSP: P06734; 1KJE.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR SMART: SM00034; CLECT.1  
 DR PROSITE: PS00615; C\_TYPE\_LLECTIN\_1; 1.  
 DR PROSITE: PS00041; C\_TYPE\_LLECTIN\_2; 1.  
 KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;

KW Calcium; Signal-anchor; Phosphorylation.  
 FT DOMAIN 1 58 CITOPLASMIC (POTENTIAL).  
 FT TRANSMEM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT DOMAIN 80 301 (POTENTIAL).  
 FT DOMAIN 169 295 EXTRACELLULAR (POTENTIAL).  
 FT DISULFID 170 181 C-TYPE LECTIN (LONG FORM).  
 FT DISULFID 198 293 BY SIMILARITY.  
 FT DISULFID 271 285 BY SIMILARITY.  
 FT CARBOHYD 97 97 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 119 119 N-LINKED (GLCNAC. . .).  
 FT CARBOHYD 165 165 N-LINKED (GLCNAC. . .).  
 FT CONFLICT 153 153 R -> A (IN REF. 1).  
 FT CONFLICT 202 202 I -> N (IN REF. 2 AND 3).  
 FT CONFLICT 260 260 C -> W (IN REF. 2 AND 3).  
 SQ SEQUENCE 301 AA; 34943 MW; 3C2315B642D71279 CRC64;  
 Query Match 21.2%; Score 141.5; DB 1; Length 301;  
 Best Local Similarity 28.1%; Pred No. 2.2e-07;  
 Matches 34; Conservative 19; Mismatches 53; Indels 15; Gaps 5;  
 OY 3 YHSEKPMNQRARRRCRDNYTDVAIQNKAIEIYDEKTLPSRSYYWIGIRKIGIWTW 62  
 DB 182 YWFSRGLTWAEADQYQOMEIAHLIVINSREGEFVVKHR--GAFHWIGLIDKDGSMKW 239  
 OY 63 V-GTNSLTFEAEWNGDGEPPN-----KKNKECVETIYIKRNKDGKWNDDACHLKAAL 116  
 DB 240 VDSIT--EYNSNFKMMAFPTQDNCQGHGEGSEDCAEIL-----SDGLWMDNCCQOVNRWA 292  
 OY 117 C 117  
 DB 293 C 293  
 RESULT 29  
 ID FCE2\_MOUSE STANDARD; PRT; 331 AA.  
 AC P20693; O61556; O61557;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Low affinity immunoglobulin epsilon FC receptor (lymphocyte IGE  
 DE receptor) (Fc-epsilon-R1) (CD23).  
 GN FCE2 OR FCE2A.  
 OS Mus musculus (Mouse).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RP MEDLINE=90017519; PubMed=2529542;  
 RA Bettler B., Hofstetter H., Rao M., Yokoyama W.M., Klichherr F.,  
 RA Conrad D.H.;  
 RT "Molecular structure and expression of the murine lymphocyte low-  
 RT affinity receptor for IGE (Fc epsilon RII)".  
 RT J. Immunol. 144:1974-1982(1990).  
 RN [3]  
 RP SEQUENCE FROM N.A. (FORMS B AND C).  
 RP STRAIN=DBA/2;  
 RX MEDLINE=94372613; PubMed=8086828;  
 RA Kondo H., Ichikawa Y., Nakamura K., Tsuchiya S.;  
 RT "Cloning of cDNAs for new subtypes of murine low-affinity Fc receptor  
 RT for IGE (Fc epsilon RII/CD23)".  
 RL Int. Arch. Allergy Immunol. 105:38-48(1994).

[4]  
RN 3D-STRUCTURE MODELING OF LECTIN DOMAIN.  
RP MEDLINE=94191542; PubMed=8142907;  
RA Padan E.A., Helm B.A.;  
RT "Modeling of the lectin-homology domains of the human and murine low-  
affinity Fc epsilon receptor (Fc epsilon RI/CD23).";  
RL Receptor 3:325-341(1993).  
CC -1- FUNCTION: THIS RECEPTOR HAS ESSENTIAL ROLES IN THE REGULATION  
OF ICE PRODUCTION AND IN THE DIFFERENTIATION OF B-CELLS (IT IS  
CC A B-CELL-SPECIFIC ANTIGEN).  
CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.  
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE  
CC PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- PFM: N- AND O-GLYCOSYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: THERE ARE TWO KINDS OF FC RECEPTORS FOR ICE, WHICH  
CC DIFFER IN BOTH STRUCTURE AND FUNCTION: HIGH AFFINITY RECEPTORS ON  
CC BASOPHILS AND MAST CELLS AND LOW AFFINITY RECEPTORS ON LYMPHOCYTES  
CC AND MONOCYTES.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; M99371; AAA74898.1; -;  
DR EMBL; M34163; AAA37603.1; -;  
DR EMBL; X64223; CAA45532.1; -;  
DR EMBL; X64224; CAA45533.1; -;  
DR PIR; A43518; A43518.  
DR PDB; 1HLJ; 3I-JAN-94.  
DR MG; MGI:95497; Fcgr2a.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00059; lectin\_c; 1.  
DR SMART: SM00034; CLECT: 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LECTIN\_2; 1.  
KW Igb-binding protein; Transmembrane; Glycoprotein; Receptor; B-cell;  
KW Repeat; Lectin; Signal-anchor; Alternative splicing; 3D-structure.  
FT DOMAIN 1 23 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 24 49 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT 50 331 (POTENTIAL).  
FT DOMAIN 185 298 EXTRACELLULAR (POTENTIAL).  
FT REPEAT 71 91 C-TYPE LECTIN (LONG FORM).  
FT REPEAT 92 112  
FT REPEAT 113 133  
FT DISULFID 183 311 BY SIMILARITY.  
FT DISULFID 186 197 BY SIMILARITY.  
FT DISULFID 214 305 BY SIMILARITY.  
FT DISULFID 282 296 BY SIMILARITY.  
FT CARBOHYD 65 65 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 114 114 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VASPLIC 1 7 MENEYS -> MNSOQ (IN ISOFORM B).  
FT VASPLIC 1 7 MENEYS -> MDTHT (IN ISOFORM C).  
SQ SEQUENCE 331 AA; 37647 MW; B8C6D5F34ACDB2 CRC64;  
  
Query Match 21.2%; Score 141.5; DB 1; Length 331;  
Best Local Similarity 30.18; Pred. No. 2.5e-07;  
Matches 37; Conservative 23; Mismatches 40; Indels 23; Gaps 6;

Oy 115 ALC 117  
Db 303 WVC 305  
  
RESULT 30  
MMGL\_RAT  
ID MMGL\_RAT STANDARD: PRT; 306 AA.  
AC P49301;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Macrophage asialoglycoprotein-binding protein (M-ASGP-Bp) (Macrophage  
DE galactose/N-acetylgalactosamine-specific lectin) (MGL).  
DE Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=90293078; PubMed=2358462;  
RA Ii M., Kurata H., Itoh N., Yamashina I., Kawasaki T.;  
RT "Molecular cloning and sequence analysis of cDNA encoding the  
RT macrophage lectin specific for galactose and N-acetylgalactosamine.";  
RL J. Biol. Chem. 265:11295-11298(1990).  
RN [2]  
RP PRELIMINARY SEQUENCE OF 9-28.  
RX MEDLINE=88339956; PubMed=3421964;  
RA Ii M., Kawasaki T., Yamashina I.;  
RT "Structural similarity between the macrophage lectin specific for  
RT galactose/N-acetylgalactosamine and the hepatic asialoglycoprotein  
RT binding protein.";  
RL Biochem. Biophys. Res. Commun. 155:720-725(1988).  
CC -1- FUNCTION: RECOGNIZES TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE  
CC UNITS.  
CC -1- SUBUNIT: HOMO-OLIGOMER.  
CC -1- SUBCELLULAR LOCATION: TYPE II membrane protein.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; J05495; AAA41216.1; -;  
DR HSSP; P06734; LRJE.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam; PF00059; lectin\_c; 1.  
DR SMART; SM00034; CLECT: 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LECTIN\_2; 1.  
KW Lectin; Glycoprotein; Transmembrane; Calcium; Signal-anchor.  
FT DOMAIN 1 37 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 38 58 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT 59 306 (POTENTIAL).  
FT DOMAIN 174 300 EXTRACELLULAR (POTENTIAL).  
FT DISULFID 175 186 C-TYPE LECTIN (LONG FORM).  
FT DISULFID 203 298 BY SIMILARITY.  
FT DISULFID 276 290 BY SIMILARITY.  
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 168 168 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 306 AA; 34242 MW; D68A5DF0B98BF13 CRC64;  
  
Query Match 20.0%; Score 133.5; DB 1; Length 306;  
Best Local Similarity 27.6%; Pred. No. 1.5e-06;  
Matches 34; Conservative 20; Mismatches 50; Indels 19; Gaps 6;

Oy 3 YHSEKPMNMGRARFCNDNTDLVAIONKAEIYLEKTLPSRSRYWIGIRKIGITW 62

```

Db 187 YWTSQSGKPEPEADKTCQLENSLVVNSLAEDNFLOTJH--GSVYTWIGLTPONGFWRW 244
QY 63 V-STN--KSLTEBAENMGDEPNPN-----KKNEDCEVEIYIKRNKDAGKNWDACHIKLA 114
Db 245 VDGTDYKCKFT----HMAPRQPDNMYGHGLGGEDCAHF-----TSDGRNNDVDCQRPYR 295
QY 115 ALC 117
Db 296 WVC 298

RESULT 31
LECG_TRIST STANDARD: PRT: 158 AA.
ID LECG_TRIST STANDARD: PRT: 158 AA.
AC 09YGP1;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Galactose-binding lectin precursor (TSL).
OS Trimeresurus stejnegeri (Chinese green tree viper).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidea;
OC Viperidae; Crotalinae; Trimeresurus.
OX NCBI_TaxID=39682;
RN [1]
RP SEQUENCE FROM N.A., SEQUENCE OF 24-53, AND MASS SPECTROMETRY.
RC TISSUE=Venom;
RX MEDLINE=99348038; PubMed=10417338;
RA "Zeng R., Xu Q., Shao X.-X., Wang K.-Y., Xia Q.-C.;
RT "Characterization and analysis of a novel glycoprotein from snake
RT venom using liquid chromatography-electrospray mass spectrometry and
RT Edman degradation."
RL Eur. J. Biochem. 266:352-358(1999).
CC -I- FUNCTION: GALACTOSE-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
CC CARBOHYDRATE STRUCTURES AND AGGLUTINATE A VARIETY OF ANIMAL CELLS
CC BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS. MAY BE A
CC CALCIUM-DEPENDENT LECTIN.
CC -I- SUBUNIT: HOMODIMER; DISULFIDE-LINKED.
CC -I- MASS SPECTROMETRY: MW=17924.2; MW_ERR=2.4; METHOD=Electrospray;
CC RANGE=24-158.
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF119097; AAD17252.1;
DR HSSP: P05451; ILIT.
DR GlycoSuiteDB: Q9YGPI;
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECF; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR Lectin; Signal; Calcium; Glycoprotein.
KW SIGNAL
FT CHAIN 1 23 GALACTOSE-BINDING LECTIN.
FT DOMAIN 24 158 C-TYPE LECTIN (LONG FORM).
FT MOD_RES 33 33 OXIDIZED.
FT DISULFID 26 37 BY SIMILARITY.

```

```

FT DISULFID 54 154 BY SIMILARITY.
FT DISULFID 61 156 BY SIMILARITY.
FT DISULFID 129 146 BY SIMILARITY.
FT DISULFID 109 109 INTERCHAIN (BY SIMILARITY).
FT CARBOHYD 28 28 N-LINKED (GLCNAC. . .) (HIGH MANNOSE).
SQ SEQUENCE 158 AA; 18635 MW; 2077BC62B7A08BF9 CRC64;

Query Match 20.0%; Score 133; DB 1; Length 158;
Best Local Similarity 30.0%; Pred. No. 8, 1e-07;
Matches 39; Conservative 17; Mismatches 46; Indels 28; Gaps 9;

QY 3 YHSEKPMNQARRCRDNY---TDLVALONKAE---LEYLEKTLPSRSY----- 49
Db 38 YKIFDEPKTWEDAEFCR-KYKPGCHLASPHRLAESIDAEYIS-----DYHKROAEV 89
QY 50 WIGI--RKIGITWVGTNKSLEEAEENMGDEPNKKKEDCEVEIYIKRNKDAGKNWD 107
Db 90 WIGLLRKKDSEW--TDRSCTDYL--NMDKNDPDHYKKEFCVELVSLGYH--RWNDQ 144
QY 108 ACHIKLAALC 117
Db 145 VCESKNSFLC 154

RESULT 32
KUCR_RAT
ID KUCR_RAT STANDARD: PRT: 550 AA.
AC P10716;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Kupffer cell receptor.
DE KCLR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 83-104.
RX MEDLINE=88227939; PubMed=2836387;
RA Hoyle G.W., Hill R.L.;
RT "Molecular cloning and sequencing of a cDNA for a carbohydrate
RT binding receptor unique to rat Kupffer cells."
RL J. Biol. Chem. 263:7487-7492(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91107689; PubMed=1846367;
RA Hoyle G.W., Hill R.L.;
RT "Structure of the gene for a carbohydrate-binding receptor unique to
RT rat Kupffer cells."
RL J. Biol. Chem. 266:1850-1857(1991).
CC -I- FUNCTION: RECEPTOR WITH AN AFFINITY FOR GALACTOSE AND FUCCOSE.
CC COULD BE INVOLVED IN ENDOCYTOSIS.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein.
CC -I- TISSUE SPECIFICITY: KUPFFER CELLS.
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J03734; AAA41472.1;
DR EMBL: M55532; AAA40892.1;
DR PIR: A28166; A28166.
DR PIR: A38674; A38674.
DR HSSP: P20693; IHLL.
DR InterPro: IPR000017; Syntaxin.

```

DR InterPro; IPR001304; lectin\_c.  
DR Pfam; PF00059; lectin\_c.1.  
DR SMART; SM00034; CLECT.1.  
DR SMART; SM00503; Synn.1.  
DR PROSITE; PS00615; C-TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LLECTIN\_2; 1.  
KW Receptor; Transmembrane; Glycoprotein; Lectin; signal-anchor;  
FT DOMAIN 1 42 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 43 69 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT DOMAIN 70 550 (POTENTIAL).  
FT DOMAIN 438 538 EXTRACELLULAR (POTENTIAL).  
FT DISULFID 440 536 C-TYPE LECTIN (SHORT FORM).  
FT DISULFID 516 528 BY SIMILARITY.  
FT CARBOHYD 87 87 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 115 115 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 132 132 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 255 255 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SO SEQUENCE 550 AA; 61104 MW; 9358A6CF4C306270 CRC64;  
  
Query Match 19.8%; Score 132; DB 1; Length 550;  
Best Local Similarity 28.9%; Pred. No. 4.3e-06;  
Matches 33; Conservative 21; Mismatches 44; Indels 16; Gaps 6;  
  
QY 3 YHSEKPMNQARRCRDNTYDLVAIONKAEIYELEKTLPPSRSYWIGIRKIG--GIW 60  
DB 424 YFSSDKSMHEENFCVSGAHNLASVTSQEQAEVLQVIT--NAVDMHIGLTDQTEGNW 481  
DB 482 RWDGTPDYVQSRFRKGGPDNMRHNGEREDCV--HLQR----MMNDMAC 528  
  
RESULT 33  
LECH\_MOUSE STANDARD; PRT; 283 AA.  
AC P34927; Q64363;  
DT 01-FEB-1994 (Rel. 28, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Asialoglycoprotein receptor 1 (Hepatic lectin 1) (MHL-1) (ASGP-R)  
DE (ASGP-R).  
GN ASGRI OR ASGR-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93176818; PubMed=8439566;  
RA Takezawa R., Shinzawa K., Watanabe Y., Akaike T.;  
RT "Determination of mouse major asialoglycoprotein receptor cDNA  
sequence.";  
RL Biochim. Biophys. Acta 1172:220-222(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BALB/C; TISSUE=Liver;  
RX MEDLINE=95047431; PubMed=7958950;  
RA Monroe R.S., Huber B.E.;  
RT "The major form of the murine asialoglycoprotein receptor: cDNA  
sequence and expression in liver, testis and epididymis.";  
RL Gene 148:237-244(1994).  
-1- FUNCTION: MEDIATES THE ENDOCYTOSIS OF PLASMA GLYCOPROTEINS TO  
WHICH THE TERMINAL SIALIC ACID RESIDUE ON THEIR COMPLEX  
CARBOHYDRATE MOTIFIES HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES  
TERMINAL GALACTOSE AND N-ACETYLGALACTOSAMINE UNITS. AFTER LIGAND  
BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND  
TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE  
DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE

CC SURFACE.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein.  
CC -1- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL  
CC CELLS.  
CC -1- MISCELLANEOUS: CALCIN I C-TYPE LECTIN FAMILY DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
CC -----  
DR EMBL; D13517; BAA02734.1; -;  
DR EMBL; 009362; AAB60441.1; -;  
DR EMBL; 008372; AAB60440.1; -;  
DR PIR; S29855; S29855.  
DR HSSP; P20693; 1HLJ.  
DR MGD; MGI:80881; Asgr1.  
DR InterPro; IPR001304; lectin\_c.  
DR Pfam; PF00059; lectin\_c.1.  
DR SMART; SM00034; CLECT.1.  
DR PROSITE; PS00615; C-TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LLECTIN\_2; 1.  
KW Lectin; Glycoprotein; Receptor; Endocytosis; Transmembrane;  
KW Calcium; signal-anchor; Phosphorylation.  
FT INT\_MET 0 38 BY SIMILARITY.  
FT DOMAIN 1 38 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 39 59 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT DOMAIN 60 283 (POTENTIAL).  
FT DOMAIN 151 277 EXTRACELLULAR (POTENTIAL).  
FT SITE 4 7 C-TYPE LECTIN (LONG FORM).  
FT DISULFID 152 163 ENDOCYTOSIS SIGNAL (POTENTIAL).  
FT DISULFID 180 275 BY SIMILARITY.  
FT DISULFID 253 267 BY SIMILARITY.  
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 150 150 I -> T (IN REF. 1).  
SO SEQUENCE 283 AA; 32472 MW; 982A5D305AAED08F CRC64;  
  
Query Match 19.4%; Score 129.5; DB 1; Length 283;  
Best Local Similarity 27.0%; Pred. No. 3.6e-06;  
Matches 33; Conservative 20; Mismatches 52; Indels 17; Gaps 6;  
  
QY 3 YHSEKPMNQARRCRDNTYDLVAIONKAEIYELEKTL--PFSRSYWIGIRKIGIWT 61  
DB 164 YFSSSVRPWTEADKYCDLENANHLVVTSDQNFLGRHMGCLN--WIGLTDONGPWK 220  
QY 62 WV-GTNKSLTEAEENWGDGEPNN-----KKNEDCEVEIYIKRNKAGKAMDACHRLKA 115  
DB 221 WVDGTD--YETGFCQWNRBQPDQPNWYGHGLGSGEDCAHF-----TTDGRNNDVYCARPRYRW 273  
QY 116 LC 117  
DB 274 VC 275  
  
RESULT 34  
PSPD\_MOUSE STANDARD; PRT; 374 AA.  
ID PSPD\_MOUSE  
AC P50404;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).  
GN SPFPD OR SPFP4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;





FT MOD\_RES 176 176 HYDROXYLATION.  
 FT CONFLICT 89 89 N -> E (IN REF. 2).  
 FT CONFLICT 164 164 K -> C (IN REF. 2).  
 SO SEQUENCE 374 AA; 37561 MM; DB2BB5E39DB4A3C CRC64;

Query Match 18.8%; Score 125; DB 1; Length 374;  
 Best Local Similarity 28.1%; Pred. No. 1.4e-05;

Matches 32; Conservative 25; Mismatches 45; Indels 12; Gaps 5;

OY 6 SEKPMNQRARFCRDNTDLVAIONKAEILEKTLFSPRSYWGIRKIG-GIMTWV 63  
 DB 269 SEEP--FEDAKEMCRQAGQILSPRSATENAVALVTVAHSAKAPLSMTDGTGEKFTY- 325

OY 64 GTNKSLTEAEANMGDEPNKKKEDCEVEIYIKRNKDGKWNDDACHLKAALC 117  
 DB 326 PTGEALV--YSNMAPGEPNNNGAENCVEITFN-----GQWMDKACGEORLVIC 372

RESULT 36

MABC\_MOUSE STANDARD; PRT: 244 AA.

AC PA1317;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Mannose-binding protein C precursor (MBP-C) (Mannan-binding protein)  
 DE (RA-reactive factor p28a subunit) (RRF/p28a).  
 GN MBL2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CBA/J; TISSUE=Liver;  
 RX MEDLINE=91302823; PubMed=1712818;  
 RA Sastiy K., Zahedi K., Lelias J.M., Whitehead A.S., Ezekowitz R.A.;  
 RT "Molecular characterization of the mouse mannose-binding proteins.  
 RT The mannose-binding protein A but not C is an acute phase reactant.";  
 RL J. Immunol. 147:692-697(1991).

RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/C;  
 RX MEDLINE=95284466; PubMed=7766991;  
 RA Sastiy R., Wang J.S., Brown D.C., Ezekowitz R.A., Tauber A.I.,  
 RT "Characterization of murine mannose-binding protein genes Mb1 and  
 RT Mb2 reveals features common to other collectin genes.";  
 RL Mamm. Genome 6:103-110(1995).

RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Kuge S., Ihara S., Watanabe E., Watanabe M., Takishima K., Suga T.,  
 RA Mamaiya G., Kawakami M.;  
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: BINDS MANNOSE AND N-ACETYLGLUCOSAMINE IN A CALCIUM-  
 CC DEPENDENT MANNER. IS CAPABLE OF HOST DEFENSE AGAINST PATHOGENS.  
 CC BY ACTIVATING THE CLASSICAL COMPLEMENT PATHWAY INDEPENDENTLY OF  
 CC THE ANTIBODY.  
 CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.  
 CC -1- SUBCELLULAR LOCATION: THIS PROTEIN IS LOCATED INTRACELLULAR, MOST  
 CC PREDOMINANTLY IN THE ROUGH ENDOPLASMIC RETICULUM AND IN THE GOLGI  
 CC APPARATUS, AND LITTLE OR NO MBP IS PRESENT IN PLASMA MEMBRANES AND  
 CC LYSOSOMES. THE MBP IS EXCLUSIVELY LOCALIZED IN THE CISTERNAL  
 CC SPACE OF THE ORGANELLES, PROBABLY AS A LOOSELY BOUND MEMBRANE  
 CC PROTEIN.

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----  
 DR EMBL; S42294; AAB19343.1; -;  
 DR EMBL; U09016; AAA82010.1; -;  
 DR EMBL; U09013; AAA82010.1; JOINED.  
 DR EMBL; U09014; AAA82010.1; JOINED.  
 DR EMBL; U09015; AAA82010.1; JOINED.  
 DR EMBL; D11440; BAA02005.1; -;  
 DR HSSP; P08661; IRDO.  
 DR MGD; MGI:96924; Mb12.  
 DR InterPro; IPR000087; Collagen.  
 DR InterPro; IPR001304; lectin\_c.  
 DR Pfam; PF00059; lectin\_c; 1.  
 DR SMART; SM00034; CLECT; 1.  
 DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE; PS50041; C-TYPE LECTIN\_2; 1.  
 KW Lectin; Hydroxylation; Liver; Glycoprotein; Mannose-binding; Membrane;

KW Calcium; Collagen; Repeat; Signal.  
 FT SIGNAL 1 18 BY SIMILARITY.  
 FT CHAIN 19 244 MANNOSE-BINDING PROTEIN C.  
 FT DOMAIN 38 96 COLLAGEN-LIKE.  
 FT DOMAIN 149 242 C-TYPE LECTIN (SHORT FORM).  
 FT MOD\_RES 43 43 HYDROXYLATION (POTENTIAL).  
 FT MOD\_RES 58 58 HYDROXYLATION (POTENTIAL).  
 FT MOD\_RES 69 69 HYDROXYLATION (POTENTIAL).  
 FT MOD\_RES 78 78 HYDROXYLATION (POTENTIAL).  
 FT MOD\_RES 81 81 HYDROXYLATION (POTENTIAL).  
 FT DISULFID 29 29 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 34 34 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 151 240 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 218 232 BY SIMILARITY.  
 FT CONFLICT 3 3 I -> L (IN REF. 1).  
 FT CONFLICT 15 15 V -> A (IN REF. 1).  
 SO SEQUENCE 244 AA; 25957 MM; 49AE84E2290DEBDA CRC64;

Query Match 18.5%; Score 123.5; DB 1; Length 244;  
 Best Local Similarity 29.6%; Pred. No. 1.3e-05;  
 Matches 34; Conservative 14; Mismatches 52; Indels 15; Gaps 4;

OY 6 SEKPMNQRARFCRDNTDLVAIONKAEILEKTLFSPRSYWGIRKIG-GIMTWV 62  
 DB 138 SVKXSLDRVAKALCEFGSVATPRNAENSAIQV---ADIAVLGITDVAVESSEFDL 194  
 OY 63 VGTNKSLTEAEANMGDEPNKKKEDCEVEIYIKRNKDGKWNDDACHLKAALC 117  
 DB 195 TGNRYRYT-----NMWDGEPNNTGDEDCVIL-----GNGKRWNDVPCSDSFLAIC 240

RESULT 37

LEC3\_MEGRO STANDARD; PRT: 162 AA.

AC P07439;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Lectin BRA-3 precursor.  
 OS Megabalanus rosa (Acorn barnacle).  
 OC Eukaryota; Metazoa; Arthropoda; Crustacea; Maxillopoda; Cirripedia;  
 OC Thoracica; Sessilia; Balanomorpha; Balanoidae; Balanidae;  
 OC Megabalanus.

OX NCBI\_TaxID=6680;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93292994; PubMed=8514190;  
 RA Takamatsu N., Takeda T., Kojima M., Helshi M., Muramoto K.,  
 RA Kamuya H., Shiba T.;

RT "Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene  
 RT structure and seasonal changes of mRNA and protein levels.";  
 RL Gene 128:251-255(1993).  
 RN [2]

```

RP SEQUENCE OF 25-162.
RA Muramoto K., Kamaya H.;
RT "The amino-acid sequence of a lectin of the acorn barnacle
RL Megabalanus rosa."
CC Biochim. Biophys. Acta 874:285-295(1986).
CC -1- FUNCTION: SUGAR-BINDING PROTEIN WHICH RECOGNIZES SPECIFIC
CC CARBOHYDRATE STRUCTURES AND AGGLUTINATE A VARIETY OF ANIMAL
CC CELLS BY BINDING TO CELL-SURFACE GLYCOPROTEINS AND GLYCOLIPIDS.
CC THIS IS A CALCIUM-DEPENDENT LECTIN. INVERTEBRATE LECTINS MAY BE
CC INVOLVED IN DEFENSE FUNCTIONS.
CC -1- SUBUNIT: HOMOTETRAMER. DISULFIDE-LINKED.
CC -1- MISCELLANEOUS: THIS LECTIN BINDS GALACTOSE.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D13302; BAA02556.1; -
DR EMBL: D13298; BAA02556.1; JOINED.
DR EMBL: D13300; BAA02556.1; JOINED.
DR EMBL: D13301; BAA02556.1; JOINED.
DR PIR: A26094; LNR3.
DR HSSP: P23806; IXX.
DR InterPro: IPR001304; lectin_c.
DR Pfam: PF00059; lectin_c; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_2; 1.
DR Lectin; Calcium; Signal.
FW SIGNAL 1 24
FT CHAIN 25 162 LECTIN BRA-3.
FT DOMAIN 25 152 C-TYPE LECTIN (LONG FORM).
FT DISULFID 26 39
FT DISULFID 56 150
FT DISULFID 125 142
FT DISULFID 157 157 INTERCHAIN (WITH C-136 IN OTHER CHAIN).
FT DISULFID 160 160 INTERCHAIN (WITH C-133 IN OTHER CHAIN).
FT VARIANT 146 146 K -> R.
SQ SEQUENCE 162 AA; 18328 MW; 8BFF149IDD1CB81 CRC64;

```

Query Match 18.3%; Score 122; DB 1; Length 162;  
 Best Local Similarity 30.6%; Pred. No. 1, 1e-05;  
 Matches 37; Conservative 21; Mismatches 47; Indels 16; Gaps 8;

```

DE (Islet cells regeneration factor) (ICRF).
GN REG1 OR REG.
OS Rattus norvegicus (Rat).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
CC NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91093273; PubMed=1985964;
RA Rouquier S., Verdier J.M., Iovanna J., Dagorn J.-C., Giorgi D.;
RT "Rat pancreatic stone protein messenger RNA. Abundant expression in
RT mature exocrine cells, regulation by food content, and sequence
RT identity with the endocrine reg transcript."
RL J. Biol. Chem. 266:786-791(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88115343; PubMed=2963000;
RA Terazono K., Yamamoto H., Takasawa S., Shiga K., Yonemura Y.,
RA Tochino Y., Okamoto H.;
RT "A novel gene activated in regenerating islets."
RL J. Biol. Chem. 263:2111-2114(1988).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93326645; PubMed=7916640;
RA Dusetli N.J., Frigerio J.M., Dagorn J.-C., Iovanna J.L.;
RT "Rapid PCR cloning and sequence determination of the rat
RT 11thosathine gene."
RL Biochim. Biophys. Acta 1174:99-102(1993).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN-WISTAR;
RA Miyashita H., Suzuki Y., Watanabe T., Uno M., Morizumi S.,
RA Yonekura H., Okamoto H.;
RT "Structure and characterization of rat Reg I gene."
RL Selkagaku 65:1082-1082(1993).
RN [5]
RP SEQUENCE OF 22-69.
RC TISSUE=pancreas;
RX MEDLINE=90031455; PubMed=2680252;
RA Adrich Z., de Caro A.M., Guidoni A.A., Wondstra M.E., Rovey M.;
RT "Characterization in rat pancreatic juice of a protein homologous to
RT the human pancreatic stone protein."
RL Comp. Biochem. Physiol. 93B:793-797(1989).
CC -1- FUNCTION: MIGHT ACT AS AN INHIBITOR OF SPONTANEOUS CALCIUM
CC CARBONATE PRECIPITATION.
CC -1- TISSUE SPECIFICITY: EXPRESSED ONLY IN REGENERATING ISLETS, BUT
CC NOT IN NORMAL PANCREATIC ISLETS, INSULINOMAS OR REGENERATING
CC LIVER.
CC -----
CC -1- SIMILARITY: TO OTHER MEMBERS OF THE C-TYPE LECTIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L07512; AAA41533.1; -
DR EMBL: M62830; AAA41974.1; -
DR EMBL: M18962; AAA42028.1; -
DR EMBL: D26164; BAA05149.1; -
DR PIR: A28351; A28351.
DR PIR: PL0147; PL0147.
DR PIR: A39081; A39081.
DR PIR: S34618; S34618.
DR HSSP: P03451; ILIT.
DR InterPro: IPR001304; lectin_c.
DR InterPro: IPR003990; pancreatitis_assoc.
DR Pfam: PF00059; lectin_c; 1.
DR PRINTS: PR01504; PNCRETTISAP.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.

```



```
FT DISULFID 2010 2026 BY SIMILARITY.
FT DISULFID 2041 2084 BY SIMILARITY.
FT DISULFID 2070 2097 BY SIMILARITY.
FT CARBOHYD 126 126 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 387 387 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 611 611 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1842 1842 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 2124 AA; 221117 MW; E30BB61593A34B1 CRC64;

Query Match 18.2%; Score 121; DB 1; Length 2124;
Best Local Similarity 31.7%; Pred. No. 0.00027;
Matches 38; Conservative 17; Mismatches 45; Indels 20; Gaps 8;

Qy 1 WT-----YHSEKPMNQARRRCRDNYDLVAIQNAIELEYEKLTPESRSYWTGI- 53
Db 1918 WTKFQGHCHYRFPDPRETWVDAERCRREGQSHLSIVTPEEQEFVNKN---AODYQWIGLN 1974

Qy 54 -RKIGITWTVWGTNKSILTEAEWNGDGPNN-KNKEDC-VEIYIKNKDAGKKNDDACH 110
Db 1975 DRTIEGDFRW-SDGSHL--QEKRRPNQDPWFATGDCVVMWHER---GEMNDVPCN 2027

RESULT 40
PGCA_MOUSE STANDARD; PRT; 2132 AA.
ID PGCA_MOUSE 064021:
AC 0612B2: 064021:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE Aggrecan core protein precursor (Cartilage-specific proteoglycan core protein) (CSPC).
DE AGC1 OR AGC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
OX NCBI_Taxid=10090;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-Cartilage;
RX MEDLINE=95104847; PubMed=7806222;
RA Walcz E., Deak F., Ethardt P., Coulter S.N., Fueleop C., Horvath P.,
RA Doege K.J., Giant T.T.;
RT "Complete coding sequence, deduced primary structure, chromosomal
RT localization, and structural analysis of murine aggrecan.";
RL Genomics 22:364-371(1994).
RN [2]
RP SEQUENCE OF 211-326 FROM N.A.
RC STRAIN-129/SV;
RX MEDLINE=95004579; PubMed=7920633;
RA Watanabe H., Kimata K., Line S., Strong D., Gao L.-Y., Kozak C.A.,
RA Yamada Y.;
RT "Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in
RT the aggrecan gene.";
RL Nat. Genet. 7:154-157(1994).
CC -1- FUNCTION: THIS PROTEOGLYCAN IS A MAJOR COMPONENT OF EXTRACELLULAR
CC MATRIX OF CARTILAGINOUS TISSUES. A MAJOR FUNCTION OF THIS PROTEIN
CC IS TO RESIST COMPRESSION IN CARTILAGE. IT BINDS AVIDLY TO
CC HYALURONIC ACID VIA AN AMINO-TERMINAL GLOBULAR REGION. MAY PLAY A
CC REGULATORY ROLE IN THE MATRIX ASSEMBLY OF THE CARTILAGE.
CC -1- SUBCELLULAR LOCATION: SECRETED; EXTRACELLULAR MATRIX (BY
CC SIMILARITY).
CC -1- DOMAIN: TWO GLOBULAR DOMAINS, G1 AND G2, COMPRISE THE AMINO
CC TERMINUS OF THE PROTEOGLYCAN, WHILE ANOTHER GLOBULAR REGION, G3,
CC MAKES UP THE COOH TERMINUS. G1 CONTAINS LINK DOMAINS AND THUS
CC CONSISTS OF THREE DISULFIDE-BONDED LOOP STRUCTURES DESIGNATED AS
CC THE A, B, B' MOTIFS. G2 IS SIMILAR TO G1. THE KERATAN SULFATE (KS)
CC AND THE CHONDROITIN SULFATE (CS) ATTACHMENT DOMAINS LIE BETWEEN G2
CC AND G3.
CC -1- PTM: CONTAINS MOSTLY CHONDROITIN SULFATE, BUT ALSO KERATAN SULFATE
```

```
CC CHAINS, N-LINKED AND O-LINKED OLIGOSACCHARIDES.
CC -1- DISEASE: DEFECTS IN AGC1 ARE THE CAUSE OF CARTILAGE MATRIX
CC DEFICIENCY (CMD), AN AUTOSOMAL RECESSIVE SYNDROME CHARACTERIZED BY
CC CLEFT PALATE, SHORT LIMBS, TAIL AND SNOUT. MUTATION IN STRAIN CMD
CC CAUSES ABSENCE OF AGGREGAN BY TRUNCATION OF THE PROTEIN (MUTATION
CC IN THE G1 DOMAIN).
CC -1- SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE V-TYPE DOMAIN.
CC -1- SIMILARITY: CONTAINS 4 LINK DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 SUSHI (SCR) DOMAIN.
CC -1- SIMILARITY: BELONGS TO THE AGGREGAN/VERSICAN PROTEOGLYCAN FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
CC EMBL: L07049; AAC37670.1; -
CC EMBL: S73722; AAB32160.1; -
CC EMBL: S73721; AAB32160.1; JOINED.
CC HSSP: P98066; ITSG.
CC MGD: MGI:99602; Agc.
CC InterPro: IPR003006; Ig_MHC.
CC InterPro: IPR003596; Ig_V.
CC InterPro: IPR000558; Link.
CC InterPro: IPR003324; SGXSG.
CC InterPro: IPR000436; Sush1_SCR_CCP.
CC InterPro: IPR001304; Lectin_C.
CC Pfam: PR00059; lectin_c.1.
CC Pfam: PR02339; sgxsg.62.
CC Pfam: PF00084; sush1.1.
CC Pfam: PF00193; xlink.4.
CC ProDom: PD000918; Link; 4.
CC SMART: SM00032; CCP; 1.
CC SMART: SM00034; CLECT; 1.
CC SMART: SM00406; IGV; 1.
CC SMART: SM00445; LINK; 4.
CC PROSITE: PS00290; IG_MHC; 1.
CC PROSITE: PS01241; LINK; 4.
CC PROSITE: PS00615; C-TYPE LECTIN.1; 1.
CC PROSITE: PS50041; C-TYPE LECTIN.2; 1.
CC GlycoProtein: Cartilage; Proteoglycan; Lectin; Signal; Sush1;
CC Repeat; Immunoglobulin domain.
CC SIGNAL 1 19
CC CHAIN 20 2132
CC DOMAIN 44 140 AGGREGAN CORE PROTEIN.
CC DOMAIN 170 247 IG-LIKE V-TYPE DOMAIN.
CC DOMAIN 268 349 LINK 1.
CC DOMAIN 504 581 LINK 2.
CC DOMAIN 602 683 LINK 3.
CC DOMAIN 1918 2044 C-TYPE LECTIN.
CC DOMAIN 2048 2106 SUSHI.
CC DOMAIN 48 140 G1-B.
CC DOMAIN 152 247 G1-B'.
CC DOMAIN 253 349 G1-B''.
CC DOMAIN 486 580 G2-B.
CC DOMAIN 587 682 G2-B'.
CC DOMAIN 805 803 KS.
CC DOMAIN 805 1231 CS-1.
CC DOMAIN 1232 1917 CS-2.
CC DOMAIN 1917 2132 G3.
CC DISULFID 51 133 BY SIMILARITY.
CC DISULFID 175 246 BY SIMILARITY.
CC DISULFID 199 220 BY SIMILARITY.
CC DISULFID 273 348 BY SIMILARITY.
CC DISULFID 297 318 BY SIMILARITY.
CC DISULFID 509 580 BY SIMILARITY.
CC DISULFID 533 554 BY SIMILARITY.
CC DISULFID 607 682 BY SIMILARITY.
```



**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:12:02 ; Search time 36.49 Seconds  
(without alignments)  
308.097 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666  
Sequence: 1 WYHYSEKPMNQRARRECR.....NKDAGKWNDDACHKLAALC 117

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 10%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	323	1 S09702	L-selectin precurs
2	666	100.0	385	1 A34015	L-selectin precurs
3	631	94.7	372	2 UC5377	L-selectin precurs
4	598	89.8	376	2 JC4892	L-selectin precurs
5	581	87.2	370	2 S2124	L-selectin precurs
6	565	84.8	372	1 A32375	L-selectin precurs
7	557	83.6	372	1 A32375	L-selectin precurs
8	455	68.3	768	2 A42755	P-selectin precurs
9	451	67.7	646	2 UN0473	P-selectin precurs
10	445	66.8	485	2 S36772	P-selectin - bovin
11	445	66.4	830	2 A30359	P-selectin precurs
12	442	66.4	768	2 I53821	P-selectin - rat
13	434	65.2	551	2 I46709	endothelial leukoc
14	425	63.8	482	2 UC5092	E-selectin - pig
15	415	62.3	610	2 A35046	E-selectin precurs
16	405	60.8	612	2 B42755	E-selectin precurs
17	161.5	24.2	404	2 A46274	HIV gp120-binding
18	157	23.6	1456	1 A36563	mannose receptor p
19	154.5	23.2	321	1 LNHUR	IgF Fc receptor II
20	153.5	23.0	304	2 UC0209	lectin, galactose/
21	150.5	22.6	311	1 LNH02A	asialoglycoprotein
22	149	22.4	207	1 LNH02A	hepatic lectin - c
23	148	22.2	1479	2 T42710	mannose receptor,
24	144	21.6	1455	1 A48925	mannose receptor p
25	143.5	21.5	301	2 S13165	asialoglycoprotein
26	142.5	21.4	309	1 S34198	IgF Fc receptor II
27	141.5	21.2	331	1 LNM5ER	IgF Fc receptor, I
28	140.5	21.1	301	1 LNM72	hepatic lectin 2 -
29	136.5	20.5	742	2 JC7595	scavenger receptor

30	133.5	20.0	306	2 A42230	lectin M-ASGP-BP p
31	132	19.8	550	2 A28166	Kupffer cell recep
32	131.5	19.7	155	2 S78774	perlecan - Halioti
33	129.5	19.4	284	2 S29855	asialoglycoprotein
34	125	18.8	374	1 A42046	surfactant protein
35	123.5	18.5	244	1 LNM5MC	mannose-binding le
36	122	18.3	162	1 LNMRC3	lectin BRA3-2 prec
37	121	18.2	165	2 A28351	pancreatic stone p
38	121	18.2	2124	2 A28452	proteoglycan core
39	121	18.2	2132	1 A5182	aggreacan precursor
40	120.5	18.1	284	1 LNM7L	hepatic lectin - r
41	119.5	17.9	301	2 A53570	collactin-43 - bov
42	119	17.9	162	1 LNMRC1	lectin BRA3-1 prec
43	119	17.9	321	2 T26152	hypothetical prote
44	118.5	17.8	2415	1 A39086	aggreacan precursor
45	118	17.7	3562	2 A47171	chondroitin sulfat

## ALIGNMENTS

```
RESULT 1
S09702
L-selectin precursor, short splice form - human
N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc
ral lymph node homing receptor Leu-8
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 26-May-2000
R:Accession: S09702
R:Camelint, D.; James, S.P.; Stamenkovic, I.; Seed, B.
Nature 342, 78-82, 1989
A>Title: Leu-8/TO1 is the human equivalent of the Mel-14 lymph node homing receptor.
A:Reference number: S06798; MUID:90044046
A:Accession: S09702
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-323 <CAM>
A:Cross-references: EMBL:X17519; NID:q3434
A>Note: this translation is not annotated in GenBank entry HSLEU8, release 111.0
C:Comment: For an alternative splice form, see PIR:A34015.
C:Genetics:
A:Gene: GDB:SEL; GDB:LNHR; LSEL; LAM1; LYAM1; LAM-1
A:Cross-references: GDB:120157; GDB:118834; OMIM:153240
A:Map position: 1q22-1q23
A:Introns: 14/3; 42/1; 171/1; 207/1; 269/1; 303/2
C:Function:
A>Description: binds with low affinity to oligosaccharides like heparan sulfate and si
ment of leukocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology
C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; Inflammat
F:1-51/Domain: signal sequence #status predicted <SIG>
F:42-168/Domain: C-type lectin homology <CCH>
F:52-323/Product: L-selectin #status predicted <MAT>
F:52-300/Domain: extracellular #status predicted <EXT>
F:173-204/Domain: EGF homology <EGF>
F:210-267/Domain: complement factor H repeat homology <FHN>
F:272-332/Domain: complement factor H repeat homology #status atypical <FN2>
F:301-318/Domain: transmembrane #status predicted <TM>
F:319-333/Domain: intracellular #status predicted <INT>
F:73,117,190,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:322/Binding site: phosphate (Ser) (covalent) #status predicted
```

```
Query Match 100.0%; Score 666; DB 1; Length 323;
Best Local Similarity 100.0%; Pred. No. 2.8e-58;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 WYHYSEKPMNQRARCRONTDLVAIONKAELEYEKTLPFSRSYWGIRKIGTW 60
DB 52 WYHYSEKPMNQRARCRONTDLVAIONKAELEYEKTLPFSRSYWGIRKIGTW 111
QY 61 TWGVTNKSLETEAEWNGDEPNKKKEDCEVETIKRKKDAGKWNDDACHKLAALC 117
```

Db 112 TWGTTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 168

RESULT 2

A34015

L-selectin precursor, long splice form - human

N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leucocyte

ral lymph node homing receptor Leu-8

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000

R:Accession: 155333; S06798; J10104; A34015; A33912

R:Ord: D.C.; Enst; T.U.; Zhou, L.J.; Rambaldi, A.; Spertini, O.; Griffin, J.; Tedder, T.

J: Biol. Chem. 265, 7760-7767, 1990

A:Title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Le

A:Reference number: 155333; MUID:90243637

A:Accession: 155333

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 14-385 <ORD>

A:Cross-references: GB:M32414; NID:g187259; PIDN:AAB60700.1; PID:g386860

R:Cammerini, D.; James, S.P.; Stamenkovic, I.; Seed, B.

Nature 342, 78-82, 1989

A:Title: Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.

A:Reference number: S06798; MUID:90044046

A:Accession: S06798

A:Molecule type: mRNA

A:Residues: 1-225, 'S', 227-385 <CAM>

A:Cross-references: EMBL:X17519; NID:g34344; PIDN:CA843536.1; PID:g4902829

A:Note: this translation is not annotated in GenBank entry HSLHE8, release 111.0

R: Tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Distche, C.M.

J. Exp. Med. 170, 123-133, 1989

A:Title: Isolation and chromosomal localization of cDNAs encoding a novel human lymphocy

otens.

A:Reference number: J10104; MUID:89310350

A:Accession: J10104

A:Molecule type: mRNA

A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TED>

A:Cross-references: GB:X16150; NID:g34448; PIDN:CA843475.1; PID:g34429

A:Note: the translated sequence in GenBank entry HSLVML, release 111.0, differs from th

R: Bowen, B.R.; Nguyen, T.; Laskey, L.A.

J. Cell Biol. 109, 421-427, 1989

A:Title: Characterization of a human homologue of the murine peripheral lymph node homin

A:Reference number: A34015; MUID:89308881

A:Accession: A34015

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 14-49, 'Y', 51-190, 'H', 192-205, 'L', 207-226, 'F', 228-385 <BOW>

A:Cross-references: GB:X16070; NID:g38092; PIDN:CA84203.1; PID:g38093

R: Siegelman, M.H.; Weissman, I.L.

Proc. Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989

A:Title: Human homologue of mouse lymph node homing receptor: evolutionary conservation

A:Reference number: A33912; MUID:89315837

A:Accession: A33912

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 14-205, 'L', 207-385 <SIE>

A:Cross-references: GB:M25280; NID:g187182; PIDN:AAC63053.1; PID:g307134

C:Comment: For an alternative splice form, see PIR:S09702.

C:Genetics:

A:Gene: GDB:SELL; GDB:LNHR; LSEL; LAML; LYAML; LAM-1

A:Cross-references: GDB:120157; GDB:118834; OMIM:153240

A:Map position: 1922-1923

A:Intons: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2

C:Function:

A:Description: binds with low affinity to oligosaccharides like heparan sulfate and sialy

ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neutroph

C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;

C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammation

F:1-51/Domain: signal sequence #status predicted <SIG>

F:42-168/Domain: C-type lectin homology <LCH>

F:52-385/Product: L-selectin #status predicted <LCH>

F:52-343/Domain: extracellular #status predicted <EXT>

F:173-204/Domain: EGF homology <EGF>

F:210-267/Domain: complement factor H repeat homology <FH1>

F:272-329/Domain: complement factor H repeat homology <FH2>

F:344-368/Domain: transmembrane #status predicted <TM>

F:369-385/Domain: intracellular #status predicted <INT>

F:73,117,190,245,259/Binding site: carbohydrate (asn) (covalent) #status predicted

F:377,380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 100.0%; Score 666; DB 1; Length 385;

Best Local Similarity 100.0%; Pred. No. 3,4e-58;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSKPMNOMORARFCNDNTDVAIONKAEIYELEKTLFSPRSYWGIRKIGTM 60

|||||

Db 52 WTYHSKPMNOMORARFCNDNTDVAIONKAEIYELEKTLFSPRSYWGIRKIGTM 111

|||||

QY 61 TWGTTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 117

|||||

Db 112 TWGTTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 168

|||||

RESULT 3

JC5377

L-selectin precursor - hamadryas baboon

C:Species: Papio hamadryas (hamadryas baboon)

C:Date: 02-Jun-1997 #sequence\_revision 12-Sep-1997 #text\_change 21-Jan-2000

R:Accession: JC5377; PC4315

R:Tsunashita, N.; Fu, H.; Berg, E.L.

Gene 181, 219-220, 1996

A:Title: PCR cloning of the cDNA encoding baboon L-selectin.

A:Reference number: JC5377; MUID:97128794

A:Accession: JC5377

A:Molecule type: mRNA

A:Residues: 1-372 <TSU1>

A:Cross-references: GB:U52074; NID:g1326148; PIDN:AAB40903.1; PID:g1326149

A:Accession: PC4315

A:Molecule type: protein

A:Residues: 37-43;142-148 <TSU2>

C:Comment: This receptor is involved in the initial adhesive interaction between lymph

sites of inflammation.

C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology

F:1-38/Domain: signal sequence #status predicted <SIG>

F:29-155/Domain: C-type lectin homology <LCH>

F:39-372/Product: L-selectin #status predicted <KAT>

F:39-157/Domain: calcium-binding #status predicted <CA>

F:160-191/Domain: EGF homology <EGF>

F:197-254/Domain: complement factor H repeat homology <FH1>

F:259-316/Domain: complement factor H repeat homology <FH2>

F:333-335/Domain: transmembrane #status predicted <TM>

F:356-372/Domain: intracellular #status predicted <INT>

Query Match 94.7%; Score 631; DB 2; Length 372;

Best Local Similarity 94.0%; Pred. No. 9.5e-55;

Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 1 WTYHSKPMNOMORARFCNDNTDVAIONKAEIYELEKTLFSPRSYWGIRKIGTM 60

|||||

Db 39 WTYHSKPMNOMORARFCNDNTDVAIONKAEIYELEKTLFSPRSYWGIRKIGTM 98

|||||

QY 61 TWGTTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 117

|||||

Db 99 TWGTTNKSILTEAEENWGDGEPNNKKNKEDCVELIYIKRNKDAGKWNDDACHIKLKAALC 155

|||||

RESULT 4

JC4892

L-selectin precursor - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 21-Jan-2000

R:Accession: JC4892

R:Qian, J.; Huang, X.; Marks, R.M.

Biochem. Biophys. Res. Commun. 225, 406-412, 1996



A:Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant protein  
A:Reference number: JC4892; M01D:96334800  
A:Accession: JC4892  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <OTA>  
A:Cross-references: GB:026535; NID:g847787; PIDN:AAA67896.1; PID:g847788  
C:Comment: This protein involved in leukocyte-endothelial adhesion; It mediates adhesion  
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;  
F:1-37/Domain: signal sequence #status predicted <Sig>  
F:29-155/Domain: C-type lectin homology <LCH>  
F:38-376/Product: L-selectin #status predicted <MAT>  
F:160-191/Domain: EGF homology <EGF>  
F:197-254/Domain: complement factor H repeat homology <FH1>  
F:259-316/Domain: complement factor H repeat homology <FH2>

Query Match	89.8%	Score 598	DB 2	Length 376
Best Local Similarity	88.0%	Pred. No. 1.8e-51		
Matches 103; Conservative	7	Mismatches 7	Indels 0	Gaps 0

QY 1 WTATYSEKPMNNQRRAPRPFCDNYDTLVAIONKAEVLEKLTLPFSRSYIWIIGRKIGTM 60  
DB 39 WTHISEKPMNNERAKRFCEYNTDTLVAIONKGEIETLEKLTLPFSRSYIWIIGRKIGTM 98  
QY 61 TWGTGKSLTEAEWNGDEEPNNKKNEDEVEIIRNNKQAGWINDACHKLAALC 117  
DB 99 TWGTGKSLTEAEWNGDEEPNNKKKTCDCEVETIKRIRSGGWINDSCCKRAALC 155

RESULT 5  
 S22124  
 L-selectin precursor - bovine  
 N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)  
 C/Species: Bos primigenius taurus (cattle)  
 C:Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
 C:Accession: S22124; A46531  
 R:Bosworth, B.T.  
 Submitted to the EMBL Data Library, October 1991  
 A:Reference number: S22123  
 A:Accession: S22124  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-370 <BOS>  
 A:Cross-references: EMBL:X62882; NID:q515; PIDN:CAA44676.1; PID:q516  
 R:Malchuck, B.; White, M.; Kurk, S.; Kishimoto, T.K.; Jutila, M.A.  
 Eur. J. Immunol. 22, 469-476, 1992  
 A:Title: Characterization of the bovine peripheral lymph node homing receptor: a lectin  
 A:Reference number: A46531; MUID:92164727  
 A:Accession: A46531  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 35-79, 'Q', 81-151, 'K' <MAL>  
 A:Note: sequence inconsistent with the nucleotide translation  
 A:Note: sequence extracted from NCBI backbone (NCBIN:85686, NCBIPI:85687)  
 C/Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;  
 F:29-115/Domain: C-type lectin homology <LCH>  
 F:160-191/Domain: EGF homology <EGF>  
 F:197-254/Domain: complement factor H repeat homology <FH1>  
 F:259-316/Domain: complement factor H repeat homology <FH2>

Query Match	87.2%;	Score 581;	DB 2;	Length 370;
Best Local Similarity	82.9%;	Pred. No. 8.3e-50;		
Matches	97;	Conservative 12;	Mismatches 8;	Indels 0;
				Gaps 0;

**Qy** 1 WTHTSYSEKPMNQRARRRRCRDYDTLVATLTKAKLELYELEKTTPRSBSYMGIRTKGTM 60  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
**Db** 39 WTHYSRRPMPWKAAPAFRENTDVLVAIIONKGELTINTKLPPSRITYIGIRKEGVG 98  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

**Qy** 61 TWGTNKSILTEAEANNGDEPNPKKNKEDCEVIYIKRNKDAGKWNDDACHKTLAALC 117  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::  
**Db** 99 TWGTNKSILTEEAKNNGAEPNNNRKSKEDCEVIYIKRNKOSGKANNDACAKATAALC 155  
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

RESULT 6  
A32375 L-selectin precursor - mouse  
N:Alternate names: Lymph cell homing receptor MEL-14; Lymphocyte surface antigen Ly-2  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jan-2000  
A:Accession: A32375; A33102; A40167; A60906  
R:Lasky, L.A.; Singer, M.S.; Yednock, T.A.; Dowbenko, D.; Fennie, C.; Rodriguez, H.;  
Cell 56, 1045-1055, 1989  
A:Title: Cloning of a lymphocyte homing receptor reveals a lectin domain.  
A:Reference number: A32375; MUID:89168433  
A:Accession: A32375  
A:Molecule type: mRNA  
A:Residues: 1-372 <LAS>  
A:Cross-references: GB:M5324; NID:g198803; PIDN:AAA39431.1; PID:g198804  
R:Siegelman, M.H.; Cheng, I.C.; Weissman, I.L.; Wakeland, E.K.  
Entrez 430000

A:Title: The mouse lymph node homing receptor is identical with the lymphocyte cell s  
A:Reference number: A35102; MUID:90263086  
A:Accession: A35102  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-372 <STE1>  
A:Cross-references: GB:M36005; NID:g199735; PIDN:AAA39722.1; PID:g199736; GB:M36058;  
R:Siegelman, M.H.; van de Rij, M.; Weissman, I.L.  
A:Title: Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing  
A:Accession number: A40167; MUID:89162048  
A:Accession: A40167  
A:Molecule type: mRNA  
A:Residues: 1-372 <STE2>  
A:Cross-references: GB:X14772; NID:g52942; PIDN:CA32880.1; PID:g52943  
A:Note: part of this sequence, including the amino end of the mature protein, was con  
R:Siegelman, M.; Bond, M.W.; Gallatin, W.M.; St.John, T.; Smith, H.T.; Fried, V.A.; W  
Science 231, 823-829, 1986  
A:Title: Cell surface molecule associated with lymphocyte homing is a ubiquitinated b  
A:Reference number: A60906; MUID:86122900  
A:Accession: A60906  
A:Molecule type: protein  
A:Residues: 'X',40,'X',42,'XXX',46,'X',48,'XXXXXX',55-56,'X',58,'XXXXXX',65,'V',67,'X'  
C:Comment: This protein is ubiquitinated.  
C:Function:  
A:Description: binds with low affinity to oligosaccharides like heparan sulfate and s  
ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neurr  
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology  
C:Keywords: cell adhesion; duplication; glycoprotein; inflammation; phosphoprotein; s  
F:29-155/Domain: signal sequence #status predicted <Sig>  
F:339-372/Product: L-selectin #status homology <LICH>  
F:39-331/Domain: extracellular #status predicted <MAT>  
F:160-191/Domain: EGF homology <EGF>  
F:197-254/Domain: complement factor H repeat homology <FHI>  
F:259-316/Domain: complement factor H repeat homology <FH2>  
F:333-355/Domain: transmembrane #status predicted <TM>  
F:356-372/Domain: intracellular #status predicted <INT>  
F:60,104,216,246,278,308,320/Binding site: carboxylate (Asn) (covalent) #status pred  
F:364/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match	84.8%;	Score 565;	DB 1;	Length 372;
Best Local Similarity	82.9%;	Pred. No. 3.2e-48;		
Matches 97;	Conservative 9;	Mismatches 11;	Indels 0;	Gaps 0;

Qy	I	MTHYISKPMNMQARRCFRCDYDVALVONAKELTELEKTLPPSRSYWYIGTRKGIM	60
		:           :           :           :           :	
Dd	39	WTMYISKPMMNDLHFCQNTDLVALÖNNRELEYELENLTPSPSYIWIWGIRKIGKM	96
		:           :           :           :           :	
Qy	61	TWVTNSKLTEEAENMGDEPNPKKNKECDEVITYIKRKKDGAKNWDACHKLKAALC	117
		:           :           :           :           :	
Dd	99	TWVTNTLTKEKENNGAGAPNNKSKECCVIITYIKRRSDGSKNWDACHCHKRKAAALC	155
		:           :           :           :           :	

```

RESULT 7
S23936
L-selectin precursor - rat
N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Dec-1993 #sequence_revision 01-Sep-1995 #text_change 20-Jun-2000
C:Accession: S23936
R:Watanabe, T.; Song, Y.; Hirayama, Y.; Tamatani, T.; Kuida, K.; Miyasaka, M.
Biochim. Biophys. Acta 1131, 321-324, 1992
A:Title: Sequence and expression of a rat CDNA for LECAM-1.
A:Reference number: S23936; MUID:92329548
A:Accession: S23936
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-372 <MAT>
A:Cross-references: GB:D10831; NID:g220801; PIDN:BA01613.1; PID:g220802
C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;
C:Keywords: transmembrane protein
F:29-155/Domain: C-type lectin homology <LCH>
F:160-191/Domain: EGF homology <EGF>
F:197-234/Domain: complement factor H repeat homology <FHL>
F:259-316/Domain: complement factor H repeat homology <FH2>

```

```

Query Match 83.6%; Score 557; DB 2; Length 372;
Best Local Similarity 81.2%; Pred. No. 2e-47;
Matches 95; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

```

```

QY 1 WTYHYSEKPMWQARRFCRDNYTDVAIONKAEIYLEKTLFPRSRYWYIGIRKIGIW 60
||||| : : : : : ||||||| : : : : : |||||||
Db 39 WTYHYSEKPMWQARRFCRDNYTDVAIONKAEIYLEKTLFPRSRYWYIGIRKIGIW 98
||||| : : : : : ||||||| : : : : : |||||||
QY 61 TWGTNKSILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 117
||||| : : : : : ||||||| : : : : : |||||||
Db 99 TWGTNKTILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 155
||||| : : : : : ||||||| : : : : : |||||||

```

```

RESULT 8
P-selectin precursor - mouse
N:Alternate names: CD62; granule membrane protein 140; PADGEM
C:Species: Mus musculus (house mouse)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 19-May-2000
C:Accession: A42755; A44899
R:Weller, A.; Isenmann, S.; Vestweber, D.
J. Biol. Chem. 267, 15176-15183, 1992
A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin.
A:Reference number: A42755; MUID:92340571
A:Accession: A42755
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-768 <MEL>
A:Cross-references: GB:M87861; NID:g200552; PIDN:AAA40008.1; PID:g200553
A:Experimental source: endothelial cells
A:Note: sequence extracted from NCBI backbone (NCBIP:109467)
R:Sanders, W.E.; Wilson, R.W.; Ballantyne, C.M.; Beaudet, A.L.
Blood 80, 795-800, 1992
A:Title: Molecular cloning and analysis of in vivo expression of murine P-selectin.
A:Reference number: A44899; MUID:92345617
A:Accession: A44899
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-723, 'E', 725-768 <SAN>
A:Cross-references: GB:M72332; NID:g193565; PIDN:AAA37712.1; PID:g193566
A:Note: sequence extracted from NCBI backbone (NCBIP:109900)
C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology; EGF
C:Keywords: cell adhesion; glycoprotein; phospholipid; phosphoprotein; transmembrane
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-768/Product: P-selectin #status predicted <MAT>
F:163-194/Domain: EGF homology <EGF>
F:200-257/Domain: complement factor H repeat homology <FHL>
F:262-319/Domain: complement factor H repeat homology <FH2>

```

```

F:324-381/Domain: complement factor H repeat homology <FH3>
F:386-443/Domain: complement factor H repeat homology <FH4>
F:448-505/Domain: complement factor H repeat homology <FH5>
F:510-567/Domain: complement factor H repeat homology <FH6>
F:580-637/Domain: complement factor H repeat homology <FH7>
F:642-699/Domain: complement factor H repeat homology <FH8>
F:710-733/Domain: complement factor H repeat homology <FH9>
F:734-768/Domain: transmembrane #status predicted <TM>
F:45,54,107,212,347,398,456,467,603,654,661,679/Binding site: carbohydrate (Asn) (cov

```

```

Query Match 68.3%; Score 455; DB 2; Length 768;
Best Local Similarity 67.5%; Pred. No. 5.2e-37;
Matches 79; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

```

```

QY 1 WTYHYSEKPMWQARRFCRDNYTDVAIONKAEIYLEKTLFPRSRYWYIGIRKIGIW 60
||||| : : : : : ||||||| : : : : : |||||||
Db 42 WTYHYSTRKAVSWNNSRVCFRRHFTDLVAIONKNEIAHLNDVIFPNSYWYIGIRKINNKW 101
||||| : : : : : ||||||| : : : : : |||||||
QY 61 TWGTNKSILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 117
||||| : : : : : ||||||| : : : : : |||||||
Db 102 TWGTNKTILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 158
||||| : : : : : ||||||| : : : : : |||||||

```

```

RESULT 9
P-selectin precursor - bovine
N:Alternate names: granule membrane protein-140
C:Species: Bos primigenius taurus (cattle)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 11-Jan-2000
C:Accession: JN0473
R:Strubel, N.A.; Nguyen, M.; Kansas, G.S.; Tedder, T.F.; Bischoff, J.
Biochem. Biophys. Res. Commun. 192, 338-344, 1993
A:Title: Isolation and characterization of a bovine cDNA encoding a functional homolo
A:Reference number: JN0473; MUID:93249394
A:Accession: JN0473
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-646 <STR>
A:Cross-references: GB:J12041; NID:g304246; PIDN:AAA30743.1; PID:g304247
C:Comment: This protein is a Ca2+ dependent receptor for myeloid cells.
C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;
C:Keywords: cell adhesion; glycoprotein; phospholipid; phosphoprotein; transmembr
F:1-41/Domain: signal sequence #status predicted <SIG>
F:42-646/Product: P-selectin #status predicted <MAT>
F:163-194/Domain: EGF homology <EGF>
F:200-257/Domain: complement factor H repeat homology <FH1>
F:262-319/Domain: complement factor H repeat homology <FH2>
F:324-381/Domain: complement factor H repeat homology <FH3>
F:386-443/Domain: complement factor H repeat homology <FH4>
F:458-515/Domain: complement factor H repeat homology <FH5>
F:520-577/Domain: complement factor H repeat homology <FH6>
F:588-611/Domain: transmembrane #status predicted <TM>
F:612-646/Domain: intracellular #status predicted <CYT>

```

```

Query Match 67.7%; Score 451; DB 2; Length 646;
Best Local Similarity 66.7%; Pred. No. 1.1e-36;
Matches 78; Conservative 10; Mismatches 29; Indels 0; Gaps 0;

```

```

QY 1 WTYHYSEKPMWQARRFCRDNYTDVAIONKAEIYLEKTLFPRSRYWYIGIRKIGIW 60
||||| : : : : : ||||||| : : : : : |||||||
Db 42 WTYHYSNKTYSWNYSRAQCYKTYTDVAIONKNEIAYLNFTIYPNSYWYIGIRKINNKW 101
||||| : : : : : ||||||| : : : : : |||||||
QY 61 TWGTNKSILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 117
||||| : : : : : ||||||| : : : : : |||||||
Db 102 TWGTNKTILTEAEWNGDGEPPNNKKNKEDCVETIYIKRNKDAGKWNDDACHLKAALC 158
||||| : : : : : ||||||| : : : : : |||||||

```

```

RESULT 10
S36772
E-selectin - bovine
C:Species: Bos primigenius taurus (cattle)

```



C:Keywords: duplication; glycoprotein; tandem repeat  
F:14-140/Domain: C-type lectin homology <LCH>  
F:182-239/Domain: complement factor H repeat homology <FHL>  
F:244-301/Domain: complement factor H repeat homology <FHD>  
F:306-364/Domain: complement factor H repeat homology <FHH>  
F:369-427/Domain: complement factor H repeat homology <FHH>  
F:433-466/Domain: complement factor H repeat homology <FHH>  
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 65.2%; Score 434; DB 2; Length 551;  
Best Local Similarity 63.2%; Pred. No. 4.3e-35;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNQRRRCRDNYTDVAIQKAELEYLEKTLFSPRSYYWIGIRKIGIW 60  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 24 WYHESANMTYDEASAYCOQRYTHLVAIQKKEIDYINSLDPSYIWIGIRKVNWV 83  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIKRKDKAGKWDACHKTKAALC 117  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 84 IWGTHKPLTEGAKWMAPEPNKKNNEDCVETIKRKDKGMNDEKSKKTLALC 140  
| | | | | : : : : : | | | | | : : | | | | | : : |

RESULT 14  
JC5092

E-selectin - pig  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 19-May-2000  
C:Accession: JC5092  
R:Winkler, H.; Brostjan, C.; Csizmadia, V.; Natarajan, G.; Anrather, J.; Bach, F.H.  
Gene 176, 67-72, 1996  
A:Title: The intron-exon structure of the porcine E-selectin-encoding gene.  
A:Reference number: JC5092; MUID:97075911  
A:Contents: endothelial cells  
A:Accession: JC5092  
A:Molecule type: DNA  
A:Residues: 1-482 <MTN>  
A:Cross-references: GB:U37521; NID:g1052974; PIDN:AAAC48680.1; PID:g1052975  
C:Comment: This protein is a member of the selectin family of adhesion molecules.  
C:Genetics:  
A:Introns: 13/1; 1421; 178/1; 237/1; 300/1; 363/1; 422/1; 459/1; 466/1  
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor  
F:13-139/Domain: C-type lectin homology <LCH>  
F:181-235/Domain: complement factor H repeat homology <FHL>  
F:240-298/Domain: complement factor H repeat homology <FHD>  
F:303-361/Domain: complement factor H repeat homology <FHH>  
F:366-420/Domain: complement factor H repeat homology <FHH>

Query Match 63.8%; Score 425; DB 2; Length 482;  
Best Local Similarity 63.2%; Pred. No. 2.9e-34;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNQRRRCRDNYTDVAIQKAELEYLEKTLFSPRSYYWIGIRKIGIW 60  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 23 WYASSTETWTFDDASAYCOQRYTHLVAIQKAELEYLSTFNVASYYWIGIRKINGTW 82  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIKRKDKAGKWDACHKTKAALC 117  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 83 TWIGTKALPLTEATWMAPEPNKKNNEDCVETIKRKDKGMNDEKSKKTLALC 139  
| | | | | : : : : : | | | | | : : | | | | | : : |

RESULT 15  
A35046

E-selectin precursor - human  
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1); LECAM-2  
C:Species: Homo sapiens (man)  
C:Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 21-Jul-2000  
C:Accession: A38615; A35046; A32606  
R:Collins, T.; Williams, A.; Johnston, G.I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone Jr., J.  
Biol. Chem. 266, 2466-2473, 1991  
A:Title: Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule 1  
A:Reference number: A38615; MUID:91115870

A:Accession: A38615  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-610 <COL>  
A:Cross-references: GB:M61893; GB:M58017; NID:g182043; PIDN:AAA52375.1; PID:g182046  
R:Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; Pasek, M.; Pittack, Proc. Natl. Acad. Sci. U.S.A. 87, 1673-1677, 1990  
A:Title: Endothelial leukocyte adhesion molecule 1: direct expression cloning and function

A:Reference number: A35046; MUID:90175359  
A:Accession: A35046  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-610 <HES>  
A:Cross-references: GB:M30640; NID:g182047; PIDN:AAA52377.1; PID:g182048  
R:Bevilacqua, M.P.; Stengelin, S.; Gimbrone Jr., M.A.; Seed, B.  
Science 243, 1160-1165, 1990

A:Title: Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophil  
A:Reference number: A32606; MUID:89162047  
A:Accession: A32606  
A:Molecule type: mRNA  
A:Residues: 1-467, 469-610 <BRV>  
A:Cross-references: GB:M24736; NID:g537523; PIDN:AAA52376.1; PID:g537524  
C:Genetics:  
A:Gene: GDB:SELE; ELAM; ESEL; ELAM1  
A:Cross-references: GDB:120612; OMIM:131210  
A:Map position: 1q22-1q25  
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor  
F:1-21/Domain: duplication; glycoprotein; tandem repeat; transmembrane protein  
F:12-138/Domain: signal sequence #status predicted <SIG>  
F:22-610/Product: endothelial leukocyte adhesion molecule 1 #status predicted <MAT>  
F:143-174/Domain: EGF homology <EGF>  
F:180-237/Domain: complement factor H repeat homology <FHL>  
F:242-299/Domain: complement factor H repeat homology <FHD>  
F:304-362/Domain: complement factor H repeat homology <FHH>  
F:367-423/Domain: complement factor H repeat homology <FHH>  
F:430-488/Domain: complement factor H repeat homology <FHH>  
F:493-547/Domain: complement factor H repeat homology <FHH>  
F:557-578/Domain: transmembrane #status predicted <TM>  
F:25,145,160,179,199,203,265,312,332,503,527/Binding site: carbohydrate (Asn) (covalent)

Query Match 62.3%; Score 415; DB 2; Length 610;  
Best Local Similarity 60.7%; Pred. No. 3.6e-33;  
Matches 71; Conservative 13; Mismatches 33; Indels 0; Gaps 0;

Oy 1 WYHYSEKPMNQRRRCRDNYTDVAIQKAELEYLEKTLFSPRSYYWIGIRKIGIW 60  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 22 WSTNSTEAMTYDEASAYCOQRYTHLVAIQKKEIDYINSLDPSYIWIGIRKVNWV 81  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Oy 61 TWVGTKSLTEAEENWGDEPNKKKEDCVETIKRKDKAGKWDACHKTKAALC 117  
| | | | | : : : : : | | | | | : : | | | | | : : |  
Db 82 VWVGTKPLTEAKWMAPEPNKKNNEDCVETIKRKDKGMNDEKSKKTLALC 138  
| | | | | : : : : : | | | | | : : | | | | | : : |

RESULT 16  
B42755

E-selectin precursor - mouse  
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)  
C:Species: Mus musculus (house mouse)  
C:Date: 04-Mar-1993 #sequence\_revision 31-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: S23174; B42755  
R:Becker-Andre, M.; van Hujsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarier, J.  
Eur. J. Biochem. 206, 401-411, 1992  
A:Title: Murine endothelial leukocyte adhesion molecule 1 is a close structural and functional homologue of human ELAM-1  
A:Reference number: S23174; MUID:92283265  
A:Accession: S23174  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-612 <BEC>  
A:Cross-references: GB:M80778; NID:g193014; PIDN:AAA37547.1; PID:g193015  
R:Weller, A.; Isenmann, S.; Vestweber, D.  
J. Biol. Chem. 267, 15176-15183, 1992



Cell 47, 657-665, 1986  
A>Title: Molecular structure of human lymphocyte receptor for immunoglobulin E.  
A:Reference number: A26067; MUID:87051737  
A:Accession: A26067  
A:Molecule type: mRNA  
A:Residues: 1-321 <KIK>  
A:Cross-references: GB:M4766; NID:9182449; PIDN:AA52435.1; PID:9182450  
A:Experimental source: EBV-transformed B lymphoblastoid cells RPMI-8866  
R:Suter, U.; Bastos, R.; Hofstetter, H.  
Nucleic Acids Res. 15, 7295-7308, 1987  
A>Title: Molecular structure of the gene and the 5'-flanking region of the human lymphoc  
A:Reference number: S03279; MUID:88015596  
A:Accession: S03279  
A>Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 157-284 <SUT>  
A:Cross-references: GB:X06049; NID:g31316  
A>Note: all exon sequences were determined but the complete sequence is not shown  
R:Matsui, M.; Nunez, R.; Sachl, Y.; Lynch, R.G.; Yodol, J.  
FEBS Lett. 335, 51-56, 1993  
A>Title: Alternative transcripts of the human CD23/Fc-epsilon-RII. A possible novel mech  
A:Reference number: S39442; MUID:94063078  
A:Accession: S39442  
A:Molecule type: DNA  
A:Residues: 1-7, 'D', 47-50 <MASI>  
A:Experimental source: splice form a'  
A:Accession: S39443  
A:Molecule type: DNA  
A:Residues: 'MNPPSQ', 47-50 <MAS2>  
A:Experimental source: splice form b'  
R:Ludin, C.; Hofstetter, H.; Sarfati, M.; Levy, C.A.; Suter, U.; Alaimo, D.; Kilchherr,  
EMBO J. 6, 109-114, 1987  
A>Title: Cloning and expression of the cDNA coding for a human lymphocyte IgE receptor.  
A:Reference number: A26164; MUID:87218454  
A:Accession: A26164  
A:Molecule type: mRNA  
A:Residues: 1-268, 'T', 270-321 <LUD>  
A:Cross-references: GB:X04772; NID:934002; PIDN:CAA28465.1; PID:g34003  
A>Note: the codon given for 268-Asn (ACC) is inconsistent with the authors' translation  
A>Note: part of this sequence, including the amino end of soluble forms of the protein,  
R:Iwata, K.; Takami, M.; Kim, C.W.; Honjo, T.; Miyoshi, T.; Tagaya, Y.; Kawabe, T.; Yodo  
Proc. Natl. Acad. Sci. U.S.A. 84, 819-823, 1987  
A>Title: Human lymphocyte Fc receptor for IgE: sequence homology of its cloned cDNA with  
A:Reference number: A26589; MUID:87118255  
A:Accession: A26589  
A:Molecule type: mRNA  
A:Residues: 1-321 <IKU>  
A:Cross-references: GB:M15059; NID:9182447; PIDN:AA52434.1; PID:9182448  
A>Note: part of this sequence, including the amino end of soluble forms of the protein,  
R:Yokota, A.; Kikutani, H.; Tanaka, T.; Sato, R.; Barsumian, E.L.; Suenura, M.; Kishimoto  
Cell 55, 611-618, 1988  
A>Title: Two species of human Fc-epsilon receptor II (Fc-epsilon-RII/CD23): tissue-speci  
A:Reference number: A31924; MUID:89028672  
A:Accession: A31924  
A:Molecule type: mRNA  
A:Residues: 'MNPPSQ', 8-14 <YOK>  
A:Cross-references: GB:M23562; NID:9182444  
A:Experimental source: splice form IIB  
R:Ueteller, M.; Sarfati, M.; Delespesse, G.  
Mol. Immunol. 26, 1105-1112, 1989  
A>Title: Mechanisms of formation of IgE-binding factors (soluble CD23)-I. Fc epsilon R I  
A:Reference number: J10132; MUID:90220658  
A:Accession: J10132  
A:Molecule type: protein  
A:Residues: 1-321 <LET>  
A:Experimental source: lymphoblastoid B cell line  
R:Rose, K.; Turcati, G.; Graber, P.; Pochon, S.; Regamey, P.O.; Jansen, K.U.; Magnenat,  
Biochem. J. 286, 819-824, 1992  
A>Title: Partial characterization of natural and recombinant human soluble CD23.  
A:Reference number: S29107; MUID:93038513  
A:Accession: S29107  
A:Molecule type: protein  
A:Residues: 152-166; 173-179; 189-212; 230-263; 268-306 <ROS>

R:Padlan, E.A.; Helm, B.A.  
submitted to the Brookhaven Protein Data Bank, June 1993  
A:Reference number: A51791; PDB:IHLI  
A:Contents: annotation; conformation by theoretical model, residues 173-285  
R:Baigorath, J.  
submitted to the Brookhaven Protein Data Bank, November 1995  
A:Reference number: A65963; PDB:IKIE  
A:Contents: annotation; conformation by theoretical model, residues 173-285  
C:Comment: The sequence of the splice form a is shown.  
C:Comment: This receptor for the Fc portion of IgE is expressed in various hematopoie  
cells.  
C:Comment: Splice form a is expressed constitutively in B-cells; b is expressed in ot  
C:Comment: Soluble IgE-binding factors are produced by proteolytic cleavage of IgE Fc  
C:Genetics:  
A:Gene: GDB:FCER2; FCE2  
A:Cross-references: GDB:118888; OMIM:151445  
A:Map position: 19p13.3-19p13.3  
A:Introns: 8/1; 46/1; 64/1; 85/1; 106/1; 127/1; 157/1; 207/3; 243/2  
C:Superfamily: IgE receptor II; C-type lectin homology  
C:Keywords: alternative splicing; B-cell; glycoprotein; immunoglobulin receptor; macr  
F:'MNPPSQ', 47-321/Product: IgE Fc receptor II, splice form b' #status predicted <SFB>  
F:'MNPPSQ', 47-321/Product: IgE Fc receptor II, splice form a' #status predicted <SPA>  
F:1-321/Product: IgE Fc receptor II, splice form a' #status predicted <INT>  
F:1-7, 'D', 47-321/Product: IgE Fc receptor II, splice form a' #status predicted <SPAI>  
F:14-20/Region: stop-transfer sequence  
F:22-45/Domain: transmembrane #status predicted <TMN>  
F:46-321/Domain: extracellular #status predicted <EXT>  
F:64-84/Region: 21-residue repeat  
F:81-321/Product: soluble IgE-binding factor (37K) #status predicted <IGE>  
F:85-105/Region: 21-residue repeat  
F:102-321/Product: soluble IgE-binding factor (33K) #status predicted <IGB>  
F:106-126/Region: 21-residue repeat  
F:125-321/Product: soluble IgE-binding factor (29K) #status predicted <IGI>  
F:148-321/Product: soluble IgE-binding factor (25-27K), long form #status experimenta  
F:150-321/Product: soluble IgE-binding factor (25-27K), short form #status experiment  
F:63/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:147-148/Cleavage site: Lys-Leu (unidentified proteinase) #status experimental  
F:149-150/Cleavage site: Arg-Met (unidentified proteinase) #status experimental  
F:191-282, 255-273/Distulfide bonds: #status experimental

Query Match 23.2%; Score 154.5; DB 1; Length 321;  
Best Local Similarity 32.2%; Pred. No. 1e-07;  
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

OY 3 YHSEKPMWQRRRCNDNYDLVAIQKAEIYLEKTPFSRSYWGIRKIG-GIW 60  
DB 175 YTFGKGTQWVHARYACDMEQQLVISHPEQDFLTGHASHTGS--WIGLRNIDLKGEF 232  
OY 61 YWGTGKSLTEFENWGDGEPNNKKNEKCEVETIYIKRKDKAGKXMDAC-HIKRAALC 117  
DB 233 INVDSH---VDYSNWAPEEPTSRSGEDCVMM-----RSGRWMDAPCDRRKLGAWVC 282

RESULT 20  
JX0209  
lectin, galactose/N-acetylgalactosamine-specific - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: JX0209; PX0009  
R:Sato, M.; Kawakami, K.; Osawa, T.; Toyoshima, S.  
J. Biochem. 111, 331-336, 1992  
A>Title: Molecular cloning and expression of cDNA encoding a galactose/N-acetylgalact  
A:Reference number: JX0209; MUID:92268032  
A:Accession: JX0209  
A:Molecule type: mRNA  
A:Residues: 1-304 <SAT>  
A:Cross-references: GB:S36676; NID:9249360; PIDN:AA822171.1; PID:g249361  
R:Oda, S.; Sato, M.; Toyoshima, S.; Osawa, T.  
J. Biochem. 104, 600-605, 1988  
A>Title: Purification and characterization of a lectin-like molecule specific for gal

```

A:Reference number: PX0009; MUID:89197865
A:Molecule type: Protein
A:Residues: 102-120;137,'X',139-151 <ODA>
C:Superfamily: hepatic lectin; C-type lectin homology
C:Keywords: glycoprotein; lectin; macrophage; transmembrane protein
F:36-61/Domain: transmembrane #status predicted <TRA>
F:173-226/Domain: C-type lectin homology <LCH>
F:74,166/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match      23.0%; Score 153.5; DB 2; Length 304;
Best Local Similarity 29.5%; Pred. No. 1.2e-07;
Matches 38; Conservative 18; Mismatches 52; Indels 21; Gaps 6;

Oy 1 WTYH-----YSEKPMNQARARPCRDNYTDLVAIQNAEIEYLEKTLPFSSRYWIGIR 54
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 177 WTEHGSCYWFSESEKSPREADKCYCLENSHLVYVNSLEQNFQNLRL--ANVYSWILTL 234

Oy 55 KIGGIWTVV-GTNSKLTLEAEKWDGDEPNN-----KKKEDCEVEIYIRKNKDAGKWDND 108
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 235 DONGFWRVWDGTD--FEKGFKNMAPLODNNFNGHGLSGEGECDAHT-----TTGGPMWDDV 287

Oy 109 CHKLKALC 117
Db 288 CQRTFRWIC 296

RESULT 21
LNH02A
asialoglycoprotein receptor H2a - human
N:Alternate names: hepatic lectin H2a
N:Contans: ASCPR: asialoglycoprotein receptor H2a; asialoglycoprotein receptor H2b
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence.revision 31-Mar-1988 #text.change 10-Dec-1999
C:Accession: A251179; A39100; B39100; I37995; A49466; B49466; S14525
R:Spies, M.; Lodish, H.F.
Proc. Natl. Acad. Sci. U.S.A. 82, 6465-6469, 1985
A:Title: Sequence of a second human asialoglycoprotein receptor: conservation of two rec
A:Reference number: A25179; MUID:86016723
A:Accession: A25179
A:Molecule type: mRNA
A:Residues: 1-311 <SP1>
A:Cross-references: GB:M11025; NID:g179080; PIDN:AAB59519.1; PID:g179081
R:Leiderkramer, G.Z.; Lodish, H.F.
J. Biol. Chem. 266, 1237-1244, 1991
A:Title: An alternatively spliced minlexon alters the subcellular fate of the human asial
A:Reference number: A39100; MUID:9109236
A:Accession: A39100
A:Molecule type: DNA; mRNA
A:Residues: 69-99 <LED>
A:Cross-references: GB:M38420; NID:g184395
A:Accession: B39100
A:Molecule type: DNA; mRNA
A:Residues: 69-81,87-99 <LE2>
A:Cross-references: GB:M38420; NID:g184395
R:Paletta, E.; Stockert, R.J.; Racevskis, J.
Hepatology 15, 395-402, 1992
A:Title: Differences in the abundance of variably spliced transcripts for the second asial
A:Reference number: I37995; MUID:92184202
A:Accession: I37995
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-23,43-81,87-311 <PAT>
A:Cross-references: EMBL:X55283; NID:934354; PIDN:CAA8997.1; PID:g34355
R:Yuk, M.H.; Lodish, H.F.
J. Cell Biol. 123, 1735-1749, 1993
A:Title: Two pathways for the degradation of the H2 subunit of the asialoglycoprotein rec
A:Reference number: A49466; MUID:94103329
A:Accession: A49466
A:Molecule type: Protein
A:Residues: 78-98 <YUK>

```

[illegible]



C:Keywords: acetylated amino end; glycoprotein; lectin; transmembrane protein  
 F:1-23/Domain: Intracellular #status predicted <INT>  
 F:24-47/Domain: transmembrane #status predicted <TRA>  
 F:48-207/Domain: extracellular #status predicted <EXT>  
 F:78-201/Domain: C-type lectin homology <LCH>  
 F:1/Modified site: acetylated amino end (Met) #status experimental  
 F:67/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 22.4%; Score 149; DB 1; Length 207;  
 Best Local Similarity 30.4%; Pred. No. 2.2e-07;  
 Matches 34; Conservative 21; Mismatches 41; Indels 16; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGI--RKIGI 60  
 Db 93 YFSLSRSMWRKAKAECEHSHLIIIDSYAKONFVMEFRTNER--FWIGLTDENGE 150  
 QY 61 TWV--GTNKSILTEAEKMGDEPN--KKNEECVEIYIKRNDAGKWNDDACHK 109  
 Db 151 QWVGGTDRSSFT----FWKEGEPNNKGFEDCAHW----TSGQWMDVYC 193

RESULT 23  
 T42710  
 mannose receptor, macrophage - mouse  
 N:Alternate names: lambda lectin; phospholipase A2 receptor  
 C:Species: Mus musculus (house mouse)  
 C>Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 09-Jun-2000  
 C:Accession: T42710  
 R:Wu, K.; Yuan, J.; Lasky, L.A.  
 J. Biol. Chem. 271, 21323-21330, 1996  
 A:Title: Characterization of a novel member of the macrophage mannose receptor type C  
 A:Reference number: 222235; MUID:9635501  
 A:Accession: T42710  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-1479 <WU>  
 A:Cross-references: EMBL:056734; NID:g1336073; PID:g1336074; PIDN:ACG52729.1  
 C:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II re  
 C:Keywords: membrane protein; receptor  
 F:186-227/Domain: fibronectin type II repeat homology <2FR>

Query Match 22.2%; Score 148; DB 2; Length 1479;  
 Best Local Similarity 30.3%; Pred. No. 2.3e-06;  
 Matches 36; Conservative 24; Mismatches 47; Indels 12; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGI--RKIGI 60  
 Db 393 YRLQAEKRSWQESKRAOCLRGDLSLSHMAELFFITKQKQVELEWIGLNDLKLQMF 452  
 QY 61 TWV--GTNKSILTEAEKMGDEPN--KKNEECVEIYIKRNDAGKWNDDACHK 117  
 Db 453 EWSGSLVSFT----HMHFEPNNFRDSDLCDVTW----GPGRWNDSPCQSLPSIC 503

RESULT 24  
 A48925  
 mannose receptor precursor, macrophage - mouse  
 C:Species: Mus musculus (house mouse)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: A48925; S21320; PC2245  
 R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowltz, R.A.  
 Blood 80, 2363-2373, 1992  
 A:Title: Characterization of the murine macrophage mannose receptor: demonstration that  
 A:Reference number: A48925; MUID:9304353  
 A:Accession: A48925  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-1455 <HAR>  
 A:Experimental source: peritoneal macrophage (NCBIP:118733)  
 A>Note: sequence extracted from NCBI Backbone  
 R:Harris, N.; Super, M.; Rits, M.; Chang, G.; Ezekowltz, R.B.

submitted to the EMBL Data Library, April 1992  
 A:Description: Characterization of the murine macrophage mannose receptor: Demonstrat  
 on.

A:Reference number: S21320  
 A:Accession: S21320  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-302; 'N', 303-1117, 'E', 1119-1455 <HA2>  
 A:Cross-references: EMBL:211974; NID:g52997; PIDN:CAA78028.1; PID:g52998  
 R:Harris, N.; Peters, L.L.; Elcher, E.M.; Rits, M.; Raspberry, D.; Elchbaum, O.G.; Su  
 Biochem. Biophys. Res. Commun. 198, 682-692, 1994  
 A:Title: The exon-intron structure and chromosomal localization of the mouse macropha  
 A:Reference number: PC2245; MUID:94128116  
 A:Accession: PC2245  
 A:Molecule type: mRNA  
 A:Residues: 35-105 <HA3>

C:Genetics: 1  
 A:Gene: Mrc1  
 A:Map position: 2  
 A:Superfamily: phospholipase A2 receptor; C-type lectin homology; fibronectin type II  
 C:Keywords: membrane protein; receptor  
 F:168-209/Domain: fibronectin type II repeat homology <2FR>  
 F:361-485/Domain: C-type lectin homology <LCH1>  
 F:943-1077/Domain: C-type lectin homology <LCH2>

Query Match 21.6%; Score 144; DB 1; Length 1455;  
 Best Local Similarity 28.1%; Pred. No. 5.7e-06;  
 Matches 34; Conservative 22; Mismatches 51; Indels 14; Gaps 4;

QY 6 SEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGI--RKIGI 61  
 Db 962 NEKKRSWQARQCKGLKGNLVSINAQEQAVTYHMRDSTFNATGNDIAHEMFMT 1021  
 QY 62 WVTGNKSLTEAEKMGDEPNKKK--EDCEIYIKRNDAGKWNDDACHK 116  
 Db 1022 -AGQGVHT--NMCKGYPGGRSSLSYEDACVVIGNSREAGTWMDDTCSKQGYI 1076  
 QY 117 C 117  
 Db 1077 C 1077

RESULT 25  
 S13165  
 asialoglycoprotein receptor - mouse  
 N:Alternate names: hepatic lectin  
 C:Species: Mus musculus (house mouse)  
 C>Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Aug-1999  
 C:Accession: S13165  
 R:Sanford, J.P.; Doyle, D.  
 Biochim. Biophys. Acta 1087, 259-261, 1990  
 A:Title: Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor ge  
 A:Reference number: S13165; MUID:91027942  
 A:Accession: S13165  
 A:Molecule type: mRNA  
 A:Residues: 1-301 <SAN>  
 A:Cross-references: EMBL:X53042; NID:g53104; PIDN:CAA37211.1; PID:g53105  
 C:Superfamily: hepatic lectin; C-type lectin homology  
 C:Keywords: glycoprotein; liver; transmembrane protein  
 F:170-293/Domain: C-type lectin homology <LCH>

Query Match 21.5%; Score 143.5; DB 2; Length 301;  
 Best Local Similarity 28.1%; Pred. No. 1.2e-06;  
 Matches 34; Conservative 22; Mismatches 50; Indels 15; Gaps 5;

QY 3 YHSEKPMNQRARPCRDNTDLVAIONKAEIYLEKTLFPRSRYWIGIRKIGI 62  
 Db 182 YWFSRDGLTWAEDQYCOLENAHLLVINSREDFVVKHR--SQFHIMIGLTDROGSKW 239  
 QY 63 V-GTNKSLTEAEKMGDEPN--KKNEECVEIYIKRNDAGKWNDDACHK 116



Db 240 VDGTD--YRSNYRNAFTQPDNMGHGGEGEDCAEIL-----SDGHWNDFCCQVNRWV 292  
Oy 117 C 117  
Db 293 C 293

RESULT 26  
S34198  
I:GE Fc receptor II, low-affinity - rat  
N:Alternate names: CD23; Lymphocyte IgE receptor  
C:Species: Rattus norvegicus (Norway rat)  
C:Dates: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 22-Jun-1999  
C:Accession: S34198  
R:Flores-Romo, L.; Sheld, J.; Humbert, Y.; Graber, P.; Aubry, J.P.; Gauchat, J.F.; Ayat  
submitted to the EMBL Data Library, June 1993  
A:Description: Inhibition of an in vivo antigen-specific IgE response by antibodies to C  
A:Reference number: S34198  
A:Accession: S34198  
A:Molecule type: mRNA  
A:Residues: 1-309 <FIO>  
A:Cross-references: EMBL:X73579; NID:9313672; PIDN:CAA51981.1; PID:9313673  
C:Superfamily: IgE receptor II; C-type lectin homology  
C:Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat; tr  
F:1-25/Domain: intracellular #status predicted <INT>  
F:14-22/Region: stop-transfer sequence  
F:24-46/Domain: transmembrane #status predicted <TM>  
F:47-309/Domain: extracellular #status predicted <EXT>  
F:126-309/Product: soluble IgE-binding factor (29K) #status predicted <IGI>  
F:149-309/Product: soluble IgE-binding factor (25-27K) #status predicted <BFI>  
F:164-283/Domain: C-type lectin homology <LCH>  
F:192-283,260-274/Disulfide bonds: #status predicted

Query Match 21.4%; Score 142.5; DB 1; Length 309;  
Best Local Similarity 30.1%; Pred. No. 1.5e-06;  
Matches 37; Conservative 23; Mismatches 40; Indels 23; Gaps 6;  
Oy 3 YHSEKPMNQRARFCRDNTDLVAIONKAEIETLEKTLPPSRSYWIGIRKIG--GIW 60  
Db 176 YFEGGSKQWIOARFACSDLOGRLVSHSQKQDPLMOHI--NKRDSWIGLQDLNMEGEF 233  
Oy 61 TW-----VGTNKSLEEAENMGDGEPPNKKNKEDCEVEIYIKRNKDGAKMNDACAK-LKA 114  
Db 234 VWDSQSPVGYGYS-----NMNPGEPNNGGCGEDCVMM-----KSGGQNDAFCSRSLDA 280  
Oy 115 ALC 117  
Db 281 WVC 283

RESULT 27  
LNMSER  
I:GE Fc receptor, low-affinity - mouse  
N:Alternate names: Blast-2; CD23; Fc-epsilon-RII; Lymphocyte IgE receptor  
C:Species: Mus musculus (house mouse)  
C:Dates: 12-Feb-1993 #sequence\_revision 28-Oct-1994 #text\_change 22-Jun-1999  
C:Accession: A43518; A33840  
R:Goldnick, S.O.; Trounstein, M.L.; Yamashita, L.C.; Kehry, M.R.; Moore, K.W.  
J. Immunol. 144, 1974-1982, 1990  
A:Title: Isolation, characterization, and expression of cDNA clones encoding the mouse F  
A:Reference number: A43518; MUID:90171598  
A:Accession: A43518  
A:Molecule type: mRNA  
A:Residues: 1-331 <GOL>  
A:Cross-references: GB:M4163; NID:9193242; PIDN:AAA37603.1; PID:9309227  
R:Bettler, B.; Hofstetter, H.; Rao, M.; Yokoyama, W.M.; Kilchherr, F.; Conrad, D.H.  
Proc. Natl. Acad. Sci. U.S.A. 86, 7566-7570, 1989  
A:Title: Molecular structure and expression of the murine lymphocyte low-affinity recept  
A:Reference number: A33840; MUID:90017519  
A:Accession: A33840  
A:Molecule type: mRNA  
A:Residues: 1-331 <BET>

A:Cross-references: GB:M99371; NID:9193245; PIDN:AAA74898.1; PID:9193246; GB:M27150  
C:Comment: This receptor for the Fc portion of IgE is expressed in various hematopoie  
f B-cells.  
C:Superfamily: IgE receptor II; C-type lectin homology  
C:Keywords: B-cell; glycoprotein; immunoglobulin receptor; macrophage; tandem repeat;  
F:1-25/Domain: intracellular #status predicted <INT>  
F:14-22/Region: stop-transfer sequence  
F:26-46/Domain: transmembrane #status predicted <TM>  
F:47-331/Domain: extracellular #status predicted <EXT>  
F:66-86/Region: 21-residue repeat  
F:87-107/Region: 21-residue repeat  
F:108-128/Region: 21-residue repeat  
F:129-149/Region: 21-residue repeat  
F:186-305/Domain: C-type lectin homology <LCH>  
F:65,114/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 21.2%; Score 141.5; DB 1; Length 331;  
Best Local Similarity 30.1%; Pred. No. 2e-06;  
Matches 37; Conservative 23; Mismatches 40; Indels 23; Gaps 6;  
Oy 3 YHSEKPMNQRARFCRDNTDLVAIONKAEIETLEKTLPPSRSYWIGIRKIG--GIW 60  
Db 198 YFEGGSKQWIOARFACSDLOGRLVSHSQKQDPLMOHI--NKRDSWIGLQDLNMEGEF 255  
Oy 61 TW-----VGTNKSLEEAENMGDGEPPNKKNKEDCEVEIYIKRNKDGAKMNDACAK-LKA 114  
Db 256 VWDSQSPVGYGYS-----NMNPGEPNNGGCGEDCVMM-----KSGGQNDAFCSRSLDA 302  
Oy 115 ALC 117  
Db 303 WVC 305

RESULT 28  
LNRT2  
I:hepatic lectin 2 - rat  
N:Alternate names: asialoglycoprotein receptor RHL-2/3 (ASGP-R2/3)  
C:Species: Rattus norvegicus (Norway rat)  
C:Dates: 31-Mar-1988 #sequence\_revision 09-Apr-1998 #text\_change 22-Jun-1999  
C:Accession: B28462; A28462; A31601; A26888; A23417  
R:Halberg, D.F.; Wager, R.E.; Farrell, D.C.; Hildreth IV, J.; Quesenberry, M.S.; Loeb  
J. Biol. Chem. 262, 9828-9838, 1987  
A:Title: Major and minor forms of the rat liver asialoglycoprotein receptor are inde  
A:Reference number: A28462; MUID:87250656  
A:Accession: B28462  
A:Molecule type: mRNA  
A:Residues: 1-301 <HAL>  
A:Cross-references: GB:J02762; NID:9205162; PIDN:AAA41522.1; PID:9205163  
A:Accession: A28462  
A:Molecule type: protein  
A:Residues: 88-96, 'X', 98-118, 'X', 120, 129-158, 177-182, 'X', 184, 'X', 186-189, 192-290, 'C',  
R:Sanford, J.P.; Elliott, R.W.; Doyle, D.  
DNA 7, 721-728, 1988  
A:Title: Asialoglycoprotein receptor genes are linked on chromosome 11 in the mouse.  
A:Reference number: A31601; MUID:89170119  
A:Accession: A31601  
A:Molecule type: mRNA  
A:Residues: 1-301 <HAN>  
A:Cross-references: GB:X07636; NID:957066; PIDN:CAA30476.1; PID:957067  
R:McPhaul, M.; Berg, P.  
Mol. Cell. Biol. 7, 1841-1847, 1987  
A:Title: Identification and characterization of cDNA clones encoding two homologous p  
A:Reference number: A26888; MUID:87257885  
A:Accession: A26888  
A:Molecule type: mRNA  
A:Residues: 1-152, 'A', 154-201, 'T', 203-259, 'C', 261-301 <MCP>  
A:Cross-references: GB:M6347; NID:9206648; PIDN:AAA42038.1; PID:9206649  
A>Note: the authors translated the codon GCA for residue 153 as Arg and ATT for resid  
J. Biol. Chem. 259, 770-778, 1984  
A:Title: Primary structure of the rat liver asialoglycoprotein receptor: structural e  
A:Reference number: A25417; MUID:84111554

A:Accession: A25417  
A:Molecule type: Protein  
A:Residues: 201-259,'C',261-281,'ND',284-301 <DRI>  
C:Comment: Calcium is required for ligand binding.  
C:Superfamily: hepatic lectin; C-type lectin homology  
C:Keywords: endocytosis; glycoprotein; lectin; liver; receptor; transmembrane protein  
F:2-60/Domain: Intracellular #status predicted <INT>  
F:61-77/Domain: transmembrane #status predicted <TRN>  
F:78-301/Domain: extracellular #status predicted <EXT>  
F:170-293/Domain: C-type lectin homology <LCH>  
F:97,119,165/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 21.1%; Score 140.5; DB 1; Length 301;  
Best Local Similarity 28.1%; Pred. No. 2.3e-06;  
Matches 34; Conservative 19; Mismatches 53; Indels 15; Gaps 5;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKALEYLEKTLFPSRSYWIIGIRKIGITW 62  
DB 182 YWFSRDGLTWEADQYCOMENAHLLVINSREDEFFVVKHR--GAFHWIGITDKDGSWKV 239  
QY 63 V-GTNSLTBEAEWNGDEPN--KKNKEDCVETIYIKRNKDGKWDACHKLAAL 116  
DB 240 VDG--EYRNFKNMATTQPDNMNGHEGSEDCAEIL-----SDGLMNDVCOQVNRMA 292  
OY 117 C 117  
DB 293 C 293

RESULT 29  
JC7595  
scavenger receptor with C-type lectin type I - human  
C:Species: Homo sapiens (man)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7595

R:Nakamura, K.; Funakoshi, H.; Miyamoto, K.; Tokunaga, F.; Nakamura, T.  
Biochem. Biophys. Res. Commun. 280, 1028-1035, 2001  
A:title: Molecular cloning and functional characterization of a human scavenger receptor  
A:Reference number: JC7595; MUID:21092718; PMID:11162630  
A:Contents: Placenta

A:Accession: JC7595  
A:Molecule type: mRNA  
A:Residues: 1-742 <NAK>  
A:Cross-references: DDBJ:AB038518  
C:Comment: This receptor, a member of the scavenger receptor family, belonging to the tyrosine kinase receptor subfamily, is important for host defense. It forms a timer and plays a role in recognizing infected cells.  
C:Genetics:

A:Gene: srcl-I  
A:Map position: 18p11.32  
C:Keywords: coiled coil; glycoprotein; transmembrane protein  
F:1-33/Domain: cytosolic (amino-terminus) #status predicted <CYT>  
F:16-19/Region: internalization signal YKRF  
F:40-56/Domain: transmembrane #status predicted <TM>  
F:57-112/Domain: extracellular #status predicted <EXT>  
F:113-333/Domain: coiled coil #status predicted <CO>  
F:369-384/Region: serine/threonine-rich #status predicted  
F:443-589/Domain: collagen-like #status predicted <COL>  
F:607-732/Domain: C-type lectin/carbohydrate recognition domain #status predicted <CR>

Query Match 20.5%; Score 136.5; DB 2; Length 742;  
Best Local Similarity 27.9%; Pred. No. 1.5e-05;  
Matches 34; Conservative 27; Mismatches 44; Indels 17; Gaps 6;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKALEYLEKTLFPSRSYWIIGIRKIGITW 60  
DB 619 YFVSVEKEIFEDAKIFCEKSSHLVINTREDOQIKKQ--VGRSHWIGLGLDSESENEW 677  
OY 61 YWV-GTNSLTBEAEWNGDEPNKK--KEDCVETIYIKRNKDGKWDACHKLAAL 115  
DB 678 KWLDTG---SPDYKNWKAGQPDNMNGHGPEDDAGLILY-----AGQANDVQCEDEVNMF 728

OY 116 LC 117  
DB 729 IC 730

RESULT 30  
A42230  
lectin M-ASGP-BP precursor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 19-Jun-1992 #sequence\_revision 19-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A42230  
R:Li, M.; Kurata, H.; Itoh, N.; Yamashina, I.; Kawasaki, T.  
J. Biol. Chem. 265, 11295-11298, 1990  
A:title: Molecular cloning and sequence analysis of cDNA encoding the macrophage lectin  
A:Reference number: A42230; MUID:90293078  
A:Accession: A42230  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-306 <LIA>  
A:Cross-references: GB:J05495; NID:9204302; PIDN:AAA41216.1; PID:9204303  
C:Superfamily: hepatic lectin; C-type lectin homology  
F:175-298/Domain: C-type lectin homology <LCH>

Query Match 20.0%; Score 133.5; DB 2; Length 306;  
Best Local Similarity 27.6%; Pred. No. 1.2e-05;  
Matches 34; Conservative 20; Mismatches 50; Indels 19; Gaps 6;

OY 3 YHSEKPMNQRARRFRCDNTDVAIONKALEYLEKTLFPSRSYWIIGIRKIGITW 62  
DB 187 YWFSQSGKRPPEADKYCOLENSNLVYVNSLAEQNFQTHM--GSVVTWIGLTDQNGPWRW 244  
OY 63 V-GTN--KSLTEAEWNGDEPN--KKNKEDCVETIYIKRNKDGKWDACHKLAAL 114  
DB 245 VDSITDEKQFT---HWAKQPDNMWGHGLGGEDECAHF-----TSDGKMWDDVCOQPYR 295  
OY 115 ALC 117  
DB 296 WVC 298

RESULT 31  
A28166  
Kupffer cell receptor - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Sep-1989 #sequence\_revision 30-Sep-1989 #text\_change 05-Nov-1999  
C:Accession: A38674; A28166  
R:Hoyle, G.W.; Hill, R.L.  
J. Biol. Chem. 266, 1850-1857, 1991  
A:title: Structure of the gene for a carbohydrate-binding receptor unique to rat Kupffer cells  
A:Reference number: A38674; MUID:91107689  
A:Accession: A38674  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-550 <HO>  
A:Cross-references: GB:M5532; NID:9203362; PIDN:AAA40892.1; PID:9203363  
R:Hoyle, G.W.; Hill, R.L.  
J. Biol. Chem. 263, 7487-7492, 1988  
A:title: Molecular cloning and sequencing of a cDNA for a carbohydrate binding receptor  
A:Reference number: A28166; MUID:86227939  
A:Accession: A28166  
A:Molecule type: mRNA  
A:Residues: 1-550 <HO>  
A:Cross-references: GB:J03734; NID:9205050; PIDN:AAA41472.1; PID:9205051  
C:Superfamily: C-type lectin homology  
F:412-536/Domain: C-type lectin homology <LCH>

Query Match 19.8%; Score 132; DB 2; Length 550;  
Best Local Similarity 28.9%; Pred. No. 3.1e-05;  
Matches 33; Conservative 21; Mismatches 44; Indels 16; Gaps 6;



A:Accession: A42574  
 A:Molecule type: mRNA  
 A:Residues: 1-244 <KUG>  
 A:Cross-references: GB:D11440; NID:g220585; PIDN:BAA02005.1; PID:g220586  
 A:Experimental source: BALB/c, liver  
 A>Note: Sequence extracted from NCBI backbone (NCBI:P110137)  
 A:Accession: C42574  
 A:Molecule type: protein  
 A:Residues: 19, 'X', 21-28, 'X', 30-32, 72-77, 'H', 79-80, 'G', 177-185, 187-189, 'H', 191-198 <KU2>  
 A>Note: source is serum of ICR mice; differences may be allotypic  
 C:Genetics:  
 A:Gene: Mbl2  
 A:Introns: 59/1; 98/1; 121/1  
 C:Superfamily: mannose-binding lectin; C-type lectin homology  
 C:Keywords: calcium; endoplasmic reticulum; Golgi apparatus; homohexamers; hydroxyproline  
 F:1-18/Domain: signal sequence #status predicted <SIG>  
 F:19-244/Product: mannose-binding lectin C #status experimental <MAT>  
 F:38-94/Region: collagen-like  
 F:124-240/Domain: C-type lectin homology <LCH>  
 F:29,34/Disulfide bonds: interchain #status predicted  
 F:69/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 18.5%; Score 123.5; DB 1; Length 244;  
 Best Local Similarity 29.6%; Pred. No. 8.8e-05;  
 Matches 34; Conservative 14; Mismatches 52; Indels 15; Gaps 4;

Oy 6 SEKPMNORARPCRDNYTDVAIONKAEIEYLEKTLFSPSYWYIGIRK---GGIWTW 62  
 Db 138 SVKMSIDRKALCSEFGSVAITRMAEENSAIDK---MDIADITDITDREGESEDL 194  
 Oy 63 VGTNKSILTEAEWNGDEPNKKKKEDCEVEIYIKRNDAGKNDACIKKALC 117  
 Db 195 TGNRVRYT---NMNDDEPNNTGDEDCVYL-----GNGKMDVPECSDFIALC 240

RESULT 36  
 LNR3  
 Lectin BRA3-2 precursor - barnacle (Megabalanus rosa)  
 C:Species: Megabalanus rosa  
 C>Date: 31-Dec-1988 #sequence\_revision 09-Sep-1994 #text\_change 16-Jul-1999  
 C:Accession: JCI504; A26094  
 R:Takamatsu, N.; Takeda, T.; Kojima, M.; Heishi, M.; Muramoto, K.; Kamiya, H.; Shiba, T.  
 Gene 128, 231-239, 1993  
 A:Title: Acorn barnacle Megabalanus rosa lectin (BRA-3): cDNA cloning, gene structure and  
 A:Reference number: JCI503; MUID:93292994  
 A:Accession: JCI504  
 A:Molecule type: mRNA  
 A:Residues: 1-162 <TAK>  
 R:Muramoto, K.; Kamiya, H.  
 Biochim. Biophys. Acta 874, 285-295, 1986  
 A:Title: The amino-acid sequence of a lectin of the acorn barnacle Megabalanus rosa.  
 A:Reference number: A26094  
 A:Accession: A26094  
 A:Molecule type: protein  
 A:Residues: 25-162 <MUT>  
 A>Note: 146 Arg was also found  
 C:Comment: This three galactose-binding lectin is isolated from the coelomic fluid.  
 C:Comment: This protein plays important roles in defense mechanisms and in development  
 C:Comment: The molecule is a tetramer of identical chains.  
 C:Superfamily: tetralectin; C-type lectin homology  
 C:Keywords: hemolymph; homotetramer; lectin  
 F:1-24/Domain: signal sequence #status predicted <SIG>  
 F:25-162/Product: lectin BRA3-2 #status experimental <MAT>  
 F:26-150/Domain: C-type lectin homology <LCH>  
 F:26-39,56-150,125-142/Disulfide bonds: #status experimental  
 F:157/Disulfide bonds: interchain (to 160) #status experimental  
 F:160/Disulfide bonds: interchain (to 157) #status experimental

Query Match 18.3%; Score 122; DB 1; Length 162;  
 Best Local Similarity 30.6%; Pred. No. 7.9e-05;

Matches 37; Conservative 21; Mismatches 47; Indels 16; Gaps 8;

Oy 3 YHYSEKPMNORARPCRDNYTD--LVAIONKAEIEYLEKTLFSPSYWYIGIRKIG--G 58  
 Db 40 YMASTYGVFRWMDAQLACQVHPGAVYLATIOSLENARISTV--SNNRMLIGLNDLDES 97  
 Oy 59 IWTWVGNKSLTEAEWNGDEPNKKKKEDCEVEIYIKRND--AGKNDACIKKAL 116  
 Db 98 HYVW--SNGEAT--DETYWSSNNPNMNMEN--ODCGVY---NYDVTGTQWDDDCNKKNFL 149  
 Oy 117 C 117  
 Db 150 C 150

RESULT 37  
 A28351  
 pancreatic stone protein precursor - rat  
 N:Alternate names: lithostathine  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 31-Mar-1990 #sequence\_revision 31-Mar-1990 #text\_change 11-May-2000  
 C:Accession: A28351; A39081; P10147; S34618  
 R:Terazono, K.; Yamamoto, H.; Takasawa, S.; Shiga, K.; Yonemura, Y.; Tochino, Y.; Oka  
 J. Biol. Chem. 263, 2111-2114, 1988  
 A:Title: A novel gene activated in regenerating islets.  
 A:Reference number: A92704; MUID:88115343  
 A:Accession: A28351  
 A:Molecule type: mRNA  
 A:Residues: 1-165 <TER>  
 A:Cross-references: GB:M18962; NID:g206604; PIDN:AAA42028.1; PID:g206605  
 R:Bouquier, S.; Verdier, J.M.; Iovanna, J.; Dagorn, J.C.; Giorzi, D.  
 J. Biol. Chem. 266, 786-791, 1991  
 A:Title: Rat pancreatic stone protein messenger RNA. Abundant expression in mature ex  
 A:Reference number: A39081; MUID:91093273  
 A:Accession: A39081  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-165 <ROU>  
 A:Cross-references: GB:M62930; GB:J05722; NID:g206462; PIDN:AAA1974.1; PID:g206463  
 R:Radich, Z.; De Caro, A.M.; Guidoni, A.A.; Woudstra, M.E.; Rovey, M.  
 Comp. Biochem. Physiol. B 93, 793-797, 1989  
 A:Title: Characterization in rat pancreatic juice of a protein homologous to the huma  
 A:Reference number: P10147; MUID:90031455  
 A:Accession: P10147  
 A:Molecule type: protein  
 A:Residues: 22-69 <ADR>  
 A:Experimental source: pancreas  
 R:Dusetli, N.J.; Frigerio, J.M.; Dagorn, J.C.; Iovanna, J.L.  
 Biochim. Biophys. Acta 1174, 99-102, 1993  
 A:Title: Rapid PCR cloning and sequence determination of the rat lithostathine gene.  
 A:Reference number: S34618; MUID:93326645  
 A:Accession: S34618  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-165 <DUS>  
 A:Cross-references: EMBL:L07512; NID:g393208; PIDN:AAA1533.1; PID:g393209  
 C:Comment: This protein is found in pancreatic calculi of mammals. A peptide bond bet  
 C:Comment: protein into an insoluble protein at a neutral pH of 5.5 to 7.5.  
 C:Genetics:  
 A:Introns: 21/1; 60/3; 106/3; 144/1  
 C:Superfamily: tetralectin; C-type lectin homology  
 C:Keywords: pyroglyutamic acid  
 F:1-21/Domain: signal sequence #status predicted <SIG>  
 F:22-165/Product: pancreatic stone protein #status predicted <MAT>  
 F:35-161/Domain: C-type lectin homology <LCH>  
 F:22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status experi

Query Match 18.2%; Score 121; DB 2; Length 165;  
 Best Local Similarity 27.6%; Pred. No. 0.0001;  
 Matches 34; Conservative 24; Mismatches 53; Indels 12; Gaps 6;

Oy 1 WTYHYSEKPMNORARPCRD-NYTDVAIONKAEIEYLEKTLFSP---RSYWYIGIR-- 54

```
Db      45 YCYFMEDHLSWAEADLCQNNNSGYLSVLSQAEGNFLASIKESGTTAAWVWIGLHDP 104
Oy      55 KIGGIWTVWVGINKSLTEAEANMGSGEPNNKKNECEVELIYIKRNDACKWNDACHKRLA 114
Db      105 KNNRRHMHSSGSLFL--YKSMWDTGYPNN-SNRGYCVS--VTSNSGYKKMRDNSCDAQLS 158
Oy      115 ALC 117
Db      159 FVC 161

RESULT  38
A28452
proteoglycan core protein precursor, cartilage - rat
N:Alternate names: aggrecan
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 21-Jan-2000
C:Accession: A92623; A23835; A28453; A28095; A28452
R:Doerge, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 262, 17757-17767, 1987
A:Title: Complete primary structure of the rat cartilage proteoglycan core protein deduced
A:Reference number: A92623; MUID:88087070
A:Accession: A92623
A:Molecule type: mRNA
A:Residues: 1-2124 <DOE>
R:Doerge, K.; Sasaki, M.; Horigan, E.; Hassell, J.R.; Yamada, Y.
J. Biol. Chem. 263, 10040a, 1988
A:Reference number: A30069
A:Contents: annotation; revision to residue 698
R:Doerge, K.; Fernandez, P.; Hassell, J.R.; Sasaki, M.; Yamada, Y.
J. Biol. Chem. 261, 8108-8111, 1986
A:Title: Partial cDNA sequence encoding a globular domain at the C terminus of the rat c
A:Reference number: A23835; MUID:86250698
A:Accession: A23835
A:Molecule type: mRNA
A:Residues: 1856-2124 <DO2>
R:Name: P.J.; Christner, J.E.; Baker, J.R.
J. Biol. Chem. 262, 17768-17778, 1987
A:Title: Cartilage proteoglycan aggregates. The link protein and proteoglycan amino-term
A:Reference number: A28453; MUID:88087071
A:Accession: A28453
A:Molecule type: protein
A:Residues: 20-37, 'W', 39-60, 'E', 62-64, 'X', 66-69, 70-83, 84, 89-148, 'L', 150-238, 'S', 240, 'A',
C:Superfamily: aggrecan, C-type lectin homology; complement factor H repeat homology; EG
C:Keywords: glycoprotein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-2124/Product: proteoglycan core protein #status predicted <MAT>
F:44-135/Domain: immunoglobulin homology <IMM>
F:170-247/Domain: link protein repeat homology <LNK1>
F:268-349/Domain: link protein repeat homology <LNK2>
F:504-581/Domain: link protein repeat homology <LNK3>
F:602-683/Domain: link protein repeat homology <LNK4>
F:1914-2034/Domain: C-type lectin homology <LCH>
F:2041-2097/Domain: complement factor H repeat homology <FHD>
F:126,229,333,387,611,667,1842/Binding site: carbohydrate (Asn) (covalent) #status predi
```

```
A55182
aggrecan precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jan-2000
C:Accession: A55182; S55329; S50207; S51355; I78532; I58123
R:Walczak, E.; Deak, F.; Erhardt, P.; Coulter, S.N.; Fueloep, C.; Horvath, P.; Doerge, K
Genomics 22, 364-371, 1994
A:Title: Complete coding sequence, deduced primary structure, chromosomal localizatio
A:Reference number: A55182; MUID:95104847
A:Accession: A55182
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: mRNA
A:Residues: 1-2132 <MAL>
A:Cross-references: GB:I07049; NID:9678541; PIDN:AAC37670.1; PID:9191772
R:Watanabe, H.; Gao, L.; Sugiyama, S.; Doerge, K.; Kimata, K.; Yamada, Y.
Biochem. J. 308, 433-440, 1995
A:Title: Mouse aggrecan, a large cartilage proteoglycan: protein sequence, gene struc
A:Reference number: S55329; MUID:95289972
A:Accession: S55329
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-567, 'G', 569-1728, 'I', 1730-2132 <MAT1>
A:Cross-references: GB:U22901; NID:9886014
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.
Biochim. Biophys. Acta 1219, 613-622, 1994
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues
A:Reference number: S50206; MUID:95035091
A:Accession: S50207
A:Molecule type: mRNA
A:Residues: 350-481, 'R', 483-506 <GLU1>
A:Cross-references: EMBL:X80279; NID:9673432
R:Glumoff, V.
submitted to the EMBL Data Library, July 1994
A:Reference number: S51355
A:Accession: S51355
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 350-383, 'CPVMSOREPMAA' <GLU2>
A:Cross-references: EMBL:X80279
R:Watanabe, H.; Kimata, K.; Line, S.; Strong, D.; Gao, L.Y.; Kozak, C.A.; Yamada, Y.
Nature Genet. 7, 154-157, 1994
A:Title: Mouse cartilage matrix deficiency (cmd) caused by a 7 bp deletion in the agg
A:Reference number: I58123; MUID:95004579
A:Accession: I78532
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 211-326 <MAT2>
A:Cross-references: GB:S73722; NID:9765215; PIDN:AAB32160.1; PID:9765216
A:Accession: I58123
A:Molecule type: translated from GB/EMBL/DBJ
A:Residues: 211-240, 'MCTRASLRMRVRSRRKRPQNSPSPRRPPTS', 'AGCGHAMPPQASSTWGRVAVTCAALAGW'
A:Cross-references: GB:S73720; NID:9765211; PIDN:AAB32159.1; PID:9765212
C:Genetics:
A:Map position: 7
A:Introns: 253/1
C:Superfamily: aggrecan, C-type lectin homology; complement factor H repeat homology;
C:Keywords: cartilage; extracellular matrix
F:1-19/Domain: signal sequence #status predicted <SIG>
F:44-135/Domain: immunoglobulin homology <IMM>
F:170-247/Domain: link protein repeat homology <LNK1>
F:268-349/Domain: link protein repeat homology <LNK2>
F:504-581/Domain: link protein repeat homology <LNK3>
F:602-683/Domain: link protein repeat homology <LNK4>
F:1922-2042/Domain: C-type lectin homology <LCH>
F:2049-2105/Domain: complement factor H repeat homology <FHD>

Query Match      18.2%; Score 121; DB 1; Length 2132;
Best Local Similarity 31.7%; Pred. No. 0.0016;
Matches 38; Conservative 17; Mismatches 45; Indels 20; Gaps 8;

Oy      1 WT-----YHSEKPMNQARRRRCRDNYTDLVAIQNAEIEYLEKTLPFSSRYWIGI- 53
Db      1918 WTKFGCHGRFPDRFTWVDERRCREQSHLSITVPEEDEFVKN--AODVQWIGLN 1974
Oy      54 -RKIGGIWTVWVGINKSLTEAEANMGSGEPNN-KKKNEC-VEIYIKRNDACKWNDACH 110
Db      1975 DRTIEGDERW-SDGHSL--QEEKRPNPQDNFATGEDCVYIMHER-----GEANDVFCN 2027

RESULT  39
Oy      1 WT-----YHSEKPMNQARRRRCRDNYTDLVAIQNAEIEYLEKTLPFSSRYWIGI- 53
```

Db 1926 WTKFQCHCYHFPDRFTWVADEFRRCRQSHLSSTIVPEDEFPVKN---AODYQWIGLN 1982  
 QY 54 -RKIGGIWTVGVNKNKSLTEAEENWGDEPN--KKNKEDC-VEIYIKRNKDGKNDACH 110  
 Db 1983 DRTIEGDFRW-SDQSHL--QFEKWRPNQPNFATGEDCVVMWHER---GEWNDVPCN 2035

## RESULT 40

LNRTL

hepatic lectin - rat

N:Alternate names: ASGP, asialoglycoprotein receptor

C:Species: Rattus norvegicus (Norway rat)

C:Date: 28-Feb-1986 #sequence\_revision 04-Dec-1986 #text\_change 22-Jun-1999

C:Accession: A92497; A94020; B94020; A54727; A03166

R:Leung, J.O.; Holland, E.C.; Drickamer, K.

J. Biol. Chem. 260, 12523-12527, 1985

A:Title: Characterization of the gene encoding the major rat liver asialoglycoprotein re

A:Reference number: A92497; MUID:86008335

A:Accession: A92497

A:Molecule type: DNA

A:Residues: 1-284 &lt;LEU&gt;

A:Cross-references: GB:K02817; NID:q206646; PIDN:AAA42037.1; PID:q206647

R:Holland, E.C.; Leung, J.O.; Drickamer, K.

Proc. Natl. Acad. Sci. U.S.A. 81, 7338-7342, 1984

A:Title: Rat liver asialoglycoprotein receptor lacks a cleavable NH-2-terminal signal se

A:Reference number: A94020; MUID:85063786

A:Accession: A94020

A:Molecule type: mRNA

A:Residues: 1-60, 'R', 62-210 &lt;HOL&gt;

A:Experimental source: clone 22; clone 1

A:Accession: B94020

A:Molecule type: mRNA

A:Residues: 92-284 &lt;HO2&gt;

A&gt;Note: clone 22 codes for a terminator at residue 210

R:Watts, C.

Biosci. Rep. 6, 527-534, 1986

A:Title: Isolation and expression of cDNA clones for a rat liver asialoglycoprotein rece

A:Reference number: A54727; MUID:87026895

A:Accession: A54727

A:Molecule type: mRNA

A:Residues: 12-284 &lt;WAT&gt;

A:Cross-references: GB:M21770; NID:q202985; PIDN:AAA40764.1; PID:q202988

A:Experimental source: liver

C:Comment: Two types of rat hepatic lectin have been identified, RHL-1 and RHL-2/3, hav

C:Comment: After removal of sialic acid monomers from the complex carbohydrate moieties

C:Comment: The unusual orientation of this protein across the membrane is postulated to

C:Genetics:

A:introns: 23/1; 62/1; 94/1; 118/1; 147/1; 197/3; 233/2

C:Superfamily: hepatic lectin; C-type lectin homology

C:Keywords: endocytosis; glycoprotein; lectin; receptor; transmembrane protein

F:2-284/Product: hepatic lectin #status predicted &lt;MAT&gt;

F:2-39/Domain: intracellular #status predicted &lt;INT&gt;

F:40-60/Domain: transmembrane #status predicted &lt;TM&gt;

F:61-284/Domain: extracellular #status predicted &lt;EXT&gt;

F:153-276/Domain: C-type lectin homology &lt;LCH&gt;

F:75,78,146/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 18.1%; Score 120.5; DB 1; Length 284;  
 Best Local Similarity 26.2%; Pred. No. 0.0002;  
 Matches 32; Conservative 20; Mismatches 53; Indels 17; Gaps 6;

QY 3 YHSEKPMNMORARRCRDNYTDVAIQKAEIETLEKTL-PSRSYVWIGIRKIGIWT 61  
 Db 165 YWSSSVKPTWTEADKYCOLENAHLVVTSTWEGRFVQOHGPIPLT--WIGLTQNGPWK 221

QY 62 WV-GTNKSLTEAEENWGDEPN--KKNKEDC-VEIYIKRNKDGKNDACHIKLAA 115  
 Db 222 WYDGTD--YETGFKNWRPGQPDWDWGHGLGGEDCAHF-----TTDGHMNDVCRRPYRW 274

QY 116 LC 117  
 Db 222 WYDGTD--YETGFKNWRPGQPDWDWGHGLGGEDCAHF-----TTDGHMNDVCRRPYRW 274

Db 275 VC 276

Search completed: September 7, 2002, 10:15:25  
 Job time: 203 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:12:52 : Search time 85.16 Seconds  
(without alignments)  
347.469 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666

Sequence: 1 WTYHYSEKPMNMOARARFCR.....NKDAGKWNDDACHKLAALC 117

Scoring table:

Gapop 10.0, Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending\_Patents\_AA\_New.\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	323	5	US-09-791-537-53485 Sequence 53485, A
2	666	100.0	341	6	US-10-211-364-1171 Sequence 1171, Ap
3	666	100.0	341	6	US-10-212-054-1328 Sequence 1328, Ap
4	666	100.0	341	6	US-10-212-778-1158 Sequence 1158, Ap
5	666	100.0	363	5	US-09-791-537-58446 Sequence 58446, A
6	666	100.0	372	5	US-09-791-537-22816 Sequence 22816, A
7	666	100.0	372	5	US-09-791-537-42657 Sequence 42657, A
8	666	100.0	372	5	US-09-791-537-69658 Sequence 69658, A
9	666	100.0	372	5	US-09-791-537-84593 Sequence 84593, A
10	666	100.0	372	5	US-09-791-537-15267 Sequence 15267, A
11	666	100.0	385	5	US-09-791-537-51391 Sequence 51391, A
12	666	100.0	385	5	US-09-791-537-53844 Sequence 53844, A
13	666	100.0	385	5	US-09-791-537-113060 Sequence 113060, A
14	666	99.4	372	5	US-09-791-537-42659 Sequence 42659, A
15	631	94.7	372	5	US-09-791-537-42655 Sequence 42655, A
16	631	94.7	372	5	US-09-791-537-42658 Sequence 42658, A
17	598	89.8	376	5	US-09-791-537-13214 Sequence 13214, A
18	581	87.2	370	5	US-09-791-537-50403 Sequence 50403, A
19	565	84.8	360	5	US-09-791-537-60503 Sequence 60503, A
20	565	84.8	372	5	US-09-791-537-37750 Sequence 37750, A
21	557	83.6	372	5	US-09-791-537-6693 Sequence 6693, Ap
22	557	83.6	372	5	US-09-791-537-81233 Sequence 81233, A
23	464	69.7	769	5	US-09-791-537-50409 Sequence 50409, A
24	455	68.3	768	5	US-09-791-537-20989 Sequence 20989, A
25	455	68.3	768	5	US-09-791-537-37753 Sequence 37753, A
26	451	67.7	646	5	US-09-791-537-84829 Sequence 84829, A

27	445	66.8	485	5	US-09-791-537-50404 Sequence 50404, A
28	445	66.8	616	5	US-09-791-537-332025 Sequence 332025, A
29	445	66.8	740	5	US-09-791-537-332023 Sequence 332023, A
30	445	66.8	740	5	US-09-791-537-34204 Sequence 34204, A
31	445	66.8	830	5	US-09-791-537-22819 Sequence 22819, A
32	445	66.8	830	5	US-09-791-537-35618 Sequence 35618, A
33	442	66.4	768	5	US-09-791-537-86038 Sequence 86038, A
34	434	65.2	551	5	US-09-791-537-44925 Sequence 44925, A
35	434	65.2	551	5	US-09-791-537-121836 Sequence 121836, A
36	434	65.2	754	5	US-09-791-537-14651 Sequence 14651, A
37	425	63.8	482	5	US-09-791-537-41672 Sequence 41672, A
38	425	63.8	482	5	US-09-791-537-111285 Sequence 111285, A
39	425	63.8	484	5	US-09-791-537-50405 Sequence 50405, A
40	420	63.1	649	5	US-09-791-537-93873 Sequence 93873, A
41	415	62.3	162	5	US-09-791-537-23269 Sequence 23269, A
42	415	62.3	196	6	US-10-143-788-916 Sequence 916, App
43	415	62.3	610	1	PCR-US02-23913-357 Sequence 357, App
44	415	62.3	610	5	US-09-791-537-22813 Sequence 22813, A
45	415	62.3	610	5	US-09-791-537-121834 Sequence 121834, A

## ALIGNMENTS

```
RESULT 1
US-09-791-537-53485
; Sequence 53485, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biotech, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53485
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53485

Query Match 100.0%; Score 666; DB 5; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.8e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNMOARARFCRDYVTDVATONKAETLEYLKTLPFSRYWIGIRKGTW 60
    |||||||
DB 52 WTYHYSEKPMNMOARARFCRDYVTDVATONKAETLEYLKTLPFSRYWIGIRKGTW 111
    |||||||
QY 61 TWVGINKSLTEAEWNGDEPNKKKEDCEVEIYIKRNKDGKWNDDACHKLAALC 117
    |||||||
DB 112 TWVGINKSLTEAEWNGDEPNKKKEDCEVEIYIKRNKDGKWNDDACHKLAALC 168
    |||||||

RESULT 2
US-10-211-364-1171
; Sequence 1171, Application US/10211364
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P1216CIN
; CURRENT APPLICATION NUMBER: US/10/211,364
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 09/760,486
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
```

```

; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1778
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1171
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-211-364-1171
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 120
QY 61 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
    |||||||
Db 121 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 3
; Sequence 1328, Application US/10212054
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P212C1N
; CURRENT APPLICATION NUMBER: US/10/212,054
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-054-1328
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 120
*QY 61 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
```

```

Db 121 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 4
US-10-212-778-1158
; Sequence 1158, Application US/10212778
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026C1N
; CURRENT APPLICATION NUMBER: US/10/212,778
; PRIOR FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/758,449
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-778-1158
```

```

Query Match          100.0%; Score 666; DB 6; Length 341;
Best Local Similarity 100.0%; Pred. No. 2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 60
    |||||||
Db 61 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 120
```

```

QY 61 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 117
    |||||||
Db 121 TWVGTNKSLTEEAKNGDGEPPNNKKNEDCVEIYIKRNKDAGKWNDDACHKRLKALC 177
```

```

RESULT 5
US-09-791-537-58446
; Sequence 58446, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58446
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-58446
```

```

Query Match          100.0%; Score 666; DB 5; Length 363;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 60
    |||||||
Db 52 WTYHYSEKPMNQRRARFCRDNTDVAIONKAEIYLEKTLPSRSYYWIGIRKIGIWM 111
```



Oy 61 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117  
Db 112 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 168

RESULT 6  
US-09-791-537-22816  
; Sequence 22816, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 22816  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-22816

Query Match 100.0%; Score 666; DB 5; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.2e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 60  
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 98

Oy 61 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117  
Db 99 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 7  
US-09-791-537-42657  
; Sequence 42657, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 42657  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Pan troglodytes  
US-09-791-537-42657

Query Match 100.0%; Score 666; DB 5; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.2e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 60  
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 98

Oy 61 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117  
Db 99 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 8  
US-09-791-537-69658  
; Sequence 69658, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 69658  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-69658

Query Match 100.0%; Score 666; DB 5; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.2e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 60  
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 98

Oy 61 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117  
Db 99 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 9  
US-09-791-537-84593  
; Sequence 84593, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biomimix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 84593  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-791-537-84593

Query Match 100.0%; Score 666; DB 5; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.2e-64;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 60  
Db 39 WTYHSEKPMWQRRRCRDNYTDLVAIQNKAETEYLEKTLPSRSYTWIGIRKIGSIW 98

Oy 61 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117  
Db 99 TWGTTNKSLTTEAEENWGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 155

RESULT 10  
US-09-791-537-152667  
; Sequence 152667, Application US/09791537  
; GENERAL INFORMATION:

```
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152667
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667
```

```
Query Match          100.0%; Score 666; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 2,2e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 60
    |||||||
Db 39 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 99 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 155
```

```
RESULT 11
US-09-791-537-51391
; Sequence 51391, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51391
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-51391
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 12
US-09-791-537-53844
; Sequence 53844, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
```

```
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53844
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53844
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 13
US-09-791-537-113060
; Sequence 113060, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113060
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-113060
```

```
Query Match          100.0%; Score 666; DB 5; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,3e-64;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 60
    |||||||
Db 52 WTYHYSEKPMNQARARFCRDNTDVAIONKAEIYELEKTLPSRSYIWGIRKIGIW 111
```

```
QY 61 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 117
    |||||||
Db 112 TWVGTKNSLTEREAEENMGDEPNKKKNEKDCVEIYIKRNKDGKWNDDACHKLKALC 168
```

```
RESULT 14
US-09-791-537-42659
; Sequence 42659, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
```

```
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42659
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pongo pygmaeus
US-09-791-537-42659
```

```
Query Match          99.4%; Score 662; DB 5; Length 372;
Best Local Similarity 99.1%; Pred. No. 6e-64;
Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    ||||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 117
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 155
```

```
RESULT 15
US-09-791-537-42655
; Sequence 42655, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42655
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-791-537-42655
```

```
Query Match          94.7%; Score 631; DB 5; Length 372;
Best Local Similarity 94.0%; Pred. No. 1.4e-60;
Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    ||||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 117
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 155
```

```
RESULT 16
US-09-791-537-42658
; Sequence 42658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42658
```

```
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Papio hamadryas
US-09-791-537-42658
```

```
Query Match          94.7%; Score 631; DB 5; Length 372;
Best Local Similarity 94.0%; Pred. No. 1.4e-60;
Matches 110; Conservative 3; Mismatches 4; Indels 0; Gaps 0;
```

```
QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    ||||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 117
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 155
```

```
RESULT 17
US-09-791-537-132144
; Sequence 132144, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 132144
; LENGTH: 376
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-791-537-132144
```

```
Query Match          89.8%; Score 598; DB 5; Length 376;
Best Local Similarity 88.0%; Pred. No. 5.9e-57;
Matches 103; Conservative 7; Mismatches 7; Indels 0; Gaps 0;
```

```
QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    ||||||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 39 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 98
```

```
QY 61 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 117
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 99 TWVGTKNSLTEREENMGDEPNKKKEDCVETIYIKRNKDAGKWDNDACHKKAALC 155
```

```
RESULT 18
US-09-791-537-50403
; Sequence 50403, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50403
; LENGTH: 370
; TYPE: PRT
; ORGANISM: Bos taurus
US-09-791-537-50403
```

1 WTHYSEKMNQNRKRCRNYTDVAIONKAELEYLEKTLPPSRSYWIGIRKGGTW 60  
 ||||| :||: ||||| ||||| ||||| : ||||| |  
 39 WTHYSEKMNENAKRKFCKHNYTDVAIONKRELEYLEKTLPPKPYWIGIRKIGKTW 98  
 ||||| :||: ||||| ||||| ||||| : ||||| |  
 61 TWYGTAKSLTEAEKWDGGEPPNKKNEKDCVEYIYIRKNKDAGKWNDACHKLAALC 117



```
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50404
; LENGTH: 485
; TYPE: PRF
; ORGANISM: Bos taurus
US-09-791-537-50404
```

```
Query Match          66.8%; Score 445; DB 5; Length 485;
Best Local Similarity 65.8%; Pred. No. 4.1e-40;
Matches 77; Conservative 14; Mismatches 26; Indels 0; Gaps 0;
```

```
OY 1 WYHYSEKPMNORARFCRDNYTDVAIONKAEIYEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 23 WYHASTEMTTEFARDYCOKTYTALVAIONQEEIYLNSTFSSPSYWIGIRKNGTW 82
```

```
OY 61 TWGCTNKSILTEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 83 TWIGTKALITNEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 139
```

```
RESULT 28
US-09-791-537-32025
; Sequence 32025, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32025
; LENGTH: 616
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32025
```

```
Query Match          66.8%; Score 445; DB 5; Length 616;
Best Local Similarity 65.8%; Pred. No. 5.5e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 1 WYHYSEKPMNORARFCRDNYTDVAIONKAEIYEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNRYTDLVAIONKNEIDYLNKVLPEYSSYWIGIRKNNKTW 101
```

```
OY 61 TWGCTNKSILTEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALITNEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 158
```

```
RESULT 29
US-09-791-537-32023
; Sequence 32023, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32023
; LENGTH: 740
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32023
```

```
Query Match          66.8%; Score 445; DB 5; Length 740;
Best Local Similarity 65.8%; Pred. No. 6.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 1 WYHYSEKPMNORARFCRDNYTDVAIONKAEIYEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNRYTDLVAIONKNEIDYLNKVLPEYSSYWIGIRKNNKTW 101
```

```
OY 61 TWGCTNKSILTEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALITNEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 158
```

```
RESULT 30
US-09-791-537-32024
; Sequence 32024, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32024
; LENGTH: 740
; TYPE: PRF
; ORGANISM: Homo sapiens
US-09-791-537-32024
```

```
Query Match          66.8%; Score 445; DB 5; Length 740;
Best Local Similarity 65.8%; Pred. No. 6.8e-40;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 1 WYHYSEKPMNORARFCRDNYTDVAIONKAEIYEKTLPEFSRYWIGIRKIGIW 60
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 42 WYHASTKAYSWNISRYCQNRYTDLVAIONKNEIDYLNKVLPEYSSYWIGIRKNNKTW 101
```

```
OY 61 TWGCTNKSILTEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 117
    ||||| : : : : : ||||| : : : : : ||||| : : : : : ||||| : : : : :
DB 102 TWGCTKALITNEAEENMGDEGPEPNKKNEKDCVEIYIKRNKDAGKWNDDACHLKAALC 158
```

```
RESULT 31
US-09-791-537-22819
; Sequence 22819, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22819
; LENGTH: 830
```

TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-22819

Query Match 66.8%; Score 445; DB 5; Length 830;  
Best Local Similarity 65.8%; Pred. No. 7.8e-40;  
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRSYWGIRKIGIM 60  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 42 WTYHSTKAYSNISRKYCQNRYTDLVAIONKNEIDYLNKVLPIYSSYIWIGIRKNNKW 101

OY 61 TWVGTNKSLTPEAEWNGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 102 TWVGTAKALTNEAEWMDNEPNKKRNEDCVEIYIKSPAPGKWDEHCLKKKHALC 158

RESULT 32  
US-09-791-537-35618  
Sequence 35618, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791,537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 35618  
LENGTH: 830  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-791-537-35618

Query Match 66.8%; Score 445; DB 5; Length 830;  
Best Local Similarity 65.8%; Pred. No. 7.8e-40;  
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRSYWGIRKIGIM 60  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 42 WTYHSTKAYSNISRKYCQNRYTDLVAIONKNEIDYLNKVLPIYSSYIWIGIRKNNKW 101

OY 61 TWVGTNKSLTPEAEWNGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 102 TWVGTAKALTNEAEWMDNEPNKKRNEDCVEIYIKSPAPGKWDEHCLKKKHALC 158

RESULT 33  
US-09-791-537-86038  
Sequence 86038, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791,537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 86038  
LENGTH: 768  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-09-791-537-86038

Query Match 66.4%; Score 442; DB 5; Length 768;  
Best Local Similarity 64.1%; Pred. No. 1.5e-39;  
Matches 75; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRSYWGIRKIGIM 60  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 42 WTYNSTKAYSNWNSAFCKRHETDLVAIONKNEIHLNDVLPYVNSYIWIGIRKINNKW 101

OY 61 TWVGTNKSLTPEAEWNGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 102 TWVGTAKLTLEAEWMDNEPNKKRNQDCVEIYIKSNAPGKWDEHCLKKKHALC 158

RESULT 34  
US-09-791-537-44925  
Sequence 44925, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791,537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 44925  
LENGTH: 551  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-09-791-537-44925

Query Match 65.2%; Score 434; DB 5; Length 551;  
Best Local Similarity 63.2%; Pred. No. 7.6e-39;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRSYWGIRKIGIM 60  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 24 WTYHSAENMTYDEASAYCQNTLHVAIONKEIDYLSILDSPSYIWIGIRKYNKW 83

OY 61 TWVGTNKSLTPEAEWNGDEPNKKKEDCVEIYIKRNKDGKWDACHKLKALC 117  
||||| : : : : : ||||||| : : : : : ||||||| :  
DB 84 IWGTAKPLTEGAKMNAPEPNKKQNEDCVEIYIKRPDGTGMWDERCSKKKHALC 140

RESULT 35  
US-09-791-537-121836  
Sequence 121836, Application US/09791537  
GENERAL INFORMATION:  
APPLICANT: Bionomix, Inc.  
APPLICANT: Debe, Derek  
TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME  
FILE REFERENCE: 261/210  
CURRENT FILING DATE: 2001-02-22  
CURRENT APPLICATION NUMBER: US/09/791,537  
NUMBER OF SEQ ID NOS: 153055  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 121836  
LENGTH: 551  
TYPE: PRT  
ORGANISM: Oryctolagus cuniculus  
US-09-791-537-121836

Query Match 65.2%; Score 434; DB 5; Length 551;  
Best Local Similarity 63.2%; Pred. No. 7.6e-39;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRSYWGIRKIGIM 60

```
Db      24 WTYHSEAKMWTDEASAYCOQRYTHLVAIONKKEIDYLSIIDSPSYWIGIRKYNW 83
      61 TWGTTNKSILTEAEENMGDEPNKKKNEDEVEIYIKRNKDAGKWDCHIKKALC 117
      84 IWGTHKPLTEBAKAKWABEPNNKKNEDCVEIYIKRKPDGTMMNDERCSKKALC 140

RESULT 36
US-09-791-537-14651
; Sequence 14651, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14651
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-14651

Query Match
Best Local Similarity 65.2%; Score 434; DB 5; Length 754;
Matches 76; Conservative 10; Mismatches 31; Indels 0; Gaps 0;

QY      1 WTYHSEKPMNMQRRPFCRDNYTDLVAIONKAEIYELEKTLPEFSRSYWIGIRKIGIW 60
      27 WTYYSTKRAVSWNYSRIEFCQKRYTDLVAIONKKEIAYLINDVYNSYWIGIRKINDKW 86
      61 TWGTTNKSILTEAEENMGDEPNKKKNEDEVEIYIKRNKDAGKWDCHIKKALC 117
      87 TWGTTKPLTEBAENKWAENEPNNKKNODCVEIYIKSLAPGKWNDPCWKRKALC 143

RESULT 37
US-09-791-537-41672
; Sequence 41672, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 41672
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Sus scrofa domestica
US-09-791-537-41672

Query Match
Best Local Similarity 63.8%; Score 425; DB 5; Length 482;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY      1 WTYHSEKPMNMQRRPFCRDNYTDLVAIONKAEIYELEKTLPEFSRSYWIGIRKIGIW 60
      23 WYSASTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWIGIRKINGTW 82
      61 TWGTTNKSILTEAEENMGDEPNKKKNEDEVEIYIKRNKDAGKWDCHIKKALC 117
      83 TWGTTKPLTEBAENKWAENEPNNKKNODCVEIYIKSLAPGKWNDPCWKRKALC 143

RESULT 38
US-09-791-537-111285
; Sequence 111285, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 111285
; LENGTH: 482
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-111285

Query Match
Best Local Similarity 63.8%; Score 425; DB 5; Length 482;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY      1 WTYHSEKPMNMQRRPFCRDNYTDLVAIONKAEIYELEKTLPEFSRSYWIGIRKIGIW 60
      23 WYSASTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWIGIRKINGTW 82
      61 TWGTTNKSILTEAEENMGDEPNKKKNEDEVEIYIKRNKDAGKWDCHIKKALC 117
      83 TWGTTKPLTEBAENKWAENEPNNKKNODCVEIYIKSLAPGKWNDPCWKRKALC 143

RESULT 39
US-09-791-537-50405
; Sequence 50405, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomimix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 50405
; LENGTH: 484
; TYPE: PRT
; ORGANISM: Sus scrofa
US-09-791-537-50405

Query Match
Best Local Similarity 63.8%; Score 425; DB 5; Length 484;
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

QY      1 WTYHSEKPMNMQRRPFCRDNYTDLVAIONKAEIYELEKTLPEFSRSYWIGIRKIGIW 60
      23 WYSASTETMTFDDASAYCOQRYTHLVAIONHAIEIYLNSTFNYSASYWIGIRKINGTW 82
      61 TWGTTNKSILTEAEENMGDEPNKKKNEDEVEIYIKRNKDAGKWDCHIKKALC 117
      83 TWGTTKPLTEBAENKWAENEPNNKKNODCVEIYIKSLAPGKWNDPCWKRKALC 143

RESULT 40
US-09-791-537-93873
```



; Sequence 93873, Application US/09791537  
; GENERAL INFORMATION:  
; APPLICANT: Biocomix, Inc.  
; APPLICANT: Debe, Derek  
; APPLICANT: Danzer, Joseph  
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: 261/210  
; CURRENT APPLICATION NUMBER: US/09/791,537  
; CURRENT FILING DATE: 2001-02-22  
; NUMBER OF SEQ ID NOS: 153055  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 93873  
; LENGTH: 649  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
; US-09-791-537-93873

Query Match 63.1%; Score 420; DB 5; Length 649;  
Best Local Similarity 65.0%; Pred. No. 3,1e-37;  
Matches 78; Conservative 11; Mismatches 25; Indels 6; Gaps 3;  
Qy 1 WTYHYSEKPMNQARARFCRDNYTDLVATONKAIEIEYLEKTLPSRSYVIGIRK--IG 57  
DB 42 WTYHYSNKTYSMNYSRAFCQKYTDLVAIONKNEIDYLNETIPYNSYVIGIRKQQIN 101  
Qy 58 GIWTWVGITNKSLTGEAENMGDGEPPNNKKKEDCVETIYIKRNKDAGKWNDDACHLKLAALC 117  
DB 102 G--TWVG--NKNLTGEAENMADNEPNNKRNNQDCVEIYIKSLSAPGKWNDEPCWKRRRALC 158

Search completed: September 7, 2002, 10:22:08  
Job time: 556 sec

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:12:32 : Search time 304.64 Seconds  
(without alignments)  
135.182 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666  
Sequence: 1 WYHYSEKPMNMQRARFCR.....NKDAGKNWDACHKIKKALC 117

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main: \*  
1: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
7: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
8: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
9: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
10: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
11: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
13: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
14: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
15: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
16: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
17: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
18: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
19: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
20: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
21: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
22: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
23: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
24: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
25: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*  
26: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	666	100.0	341	21	US-09-758-449-1158
2	666	100.0	341	21	US-09-760-443-1328
3	666	100.0	372	1	PCT-US01-26675-3
4	666	100.0	372	15	US-09-119-209-2
5	666	100.0	382	21	US-09-760-475-2123
6	666	100.0	385	1	PCT-US94-03970-2
7	666	100.0	385	1	PCT-US94-00909-2

8	666	100.0	385	4	US-08-008-459-2	Sequence 2, Appl1
9	666	100.0 <td>385</td> <td>7</td> <td>US-08-340-539-2</td> <td>Sequence 2, Appl1</td>	385	7	US-08-340-539-2	Sequence 2, Appl1
10	666	100.0 <td>385</td> <td>8</td> <td>US-08-410-569-2</td> <td>Sequence 2, Appl1</td>	385	8	US-08-410-569-2	Sequence 2, Appl1
11	666	100.0 <td>1078</td> <td>26</td> <td>US-60-207-315-428</td> <td>Sequence 428, App</td>	1078	26	US-60-207-315-428	Sequence 428, App
12	666	100.0 <td>1078</td> <td>26</td> <td>US-60-212-659-523</td> <td>Sequence 523, App</td>	1078	26	US-60-212-659-523	Sequence 523, App
13	666	100.0 <td>1078</td> <td>26</td> <td>US-60-230-435-1751</td> <td>Sequence 1751, App</td>	1078	26	US-60-230-435-1751	Sequence 1751, App
14	654	98.2	119	26	US-60-160-189-6687	Sequence 8687, App
15	654	98.2	119	26	US-60-169-867-8823	Sequence 5823, App
16	633	95.0	119	26	US-60-160-203-5003	Sequence 5003, App
17	633	95.0	119	26	US-60-169-840-5716	Sequence 6716, App
18	565	84.8	372	15	US-09-119-209-4	Sequence 4, Appl1
19	539	80.9	116	26	US-60-160-189-10011	Sequence 10011, A
20	539	80.9	116	26	US-60-160-203-6200	Sequence 6200, App
21	539	80.9	116	26	US-60-169-840-9326	Sequence 9326, App
22	539	80.9	116	26	US-60-169-867-7998	Sequence 7998, App
23	445	66.8	700	26	US-60-207-315-467	Sequence 467, App
24	445	66.8	830	1	PCT-US94-09395-4	Sequence 4, Appl1
25	445	66.8	830	8	US-08-449-6878-4	Sequence 4, Appl1
26	445	66.8	830	24	US-10-020-141-10	Sequence 10, Appl
27	440.5	66.1	129	26	US-60-196-718-4236	Sequence 4236, App
28	440.5	66.1	133	26	US-60-196-718-4237	Sequence 4237, App
29	439.5	66.0	128	26	US-60-195-053-1909	Sequence 1909, App
30	439.5	66.0	129	26	US-60-195-053-1908	Sequence 1908, App
31	439.5	66.0	138	26	US-60-196-718-4238	Sequence 4238, App
32	415	62.3	134	26	US-60-192-729-3226	Sequence 3226, App
33	415	62.3	196	21	US-09-760-498-916	Sequence 916, App
34	415	62.3	610	1	PCT-US99-28965-19	Sequence 19, Appl
35	415	62.3	610	10	US-08-657-753-2	Sequence 2, Appl1
36	415	62.3	610	11	US-08-770-435-3	Sequence 3, Appl1
37	415	62.3	610	16	US-09-266-091-2	Sequence 2, Appl1
38	415	62.3	610	16	US-09-266-091A-2	Sequence 2, Appl1
39	415	62.3	610	21	US-09-784-356-122	Sequence 122, App
40	415	62.3	610	22	US-09-802-640-36	Sequence 36, Appl
41	415	62.3	610	22	US-09-857-670-19	Sequence 19, Appl
42	415	62.3	610	24	US-10-021-660-122	Sequence 122, App
43	404	60.7	112	26	US-60-160-203-3503	Sequence 3503, App
44	403.5	60.6	130	26	US-60-196-174-903	Sequence 903, App
45	391	58.7	104	26	US-60-160-189-5792	Sequence 5792, App

#### ALIGNMENTS

RESULT 1  
US-09-758-449-1158  
Sequence 1158, Application US/09758449  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PM026  
CURRENT FILING DATE: 2001-01-11  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
NUMBER OF SEQ ID NOS: 1478  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1158  
LENGTH: 341  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (215)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-758-449-1158

Query Match 100.0%; Score 666; DB 21; Length 341;  
Best Local Similarity 100.0%; Pred. No. 9.4e-68;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 60  
|||||  
DB 61 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 120  
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 117  
|||||  
DB 121 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 177

## RESULT 2

US-09-760-443-1328  
; Sequence 1328, Application US/09760443  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: P1212  
; CURRENT APPLICATION NUMBER: US/09/760,443  
; CURRENT FILING DATE: 2001-01-16  
; Prior application data removed - refer to PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 2164  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1328  
; LENGTH: 341  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (215)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-760-443-1328

Query Match 100.0%; Score 666; DB 21; Length 341;  
Best Local Similarity 100.0%; Pred. No. 9.4e-68;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 60  
|||||  
DB 61 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 120  
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 117  
|||||  
DB 121 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 177

RESULT 3  
PCT-US01-26675-3  
; Sequence 3, Application PC/TUS0126675  
; GENERAL INFORMATION:  
; APPLICANT: Genaisance Pharmaceuticals, Inc.  
; APPLICANT: Anastasio, Alison E  
; APPLICANT: Bieglecki, Karyn M  
; APPLICANT: Kliem, Stefanie E  
; APPLICANT: Koshy, Beena  
; APPLICANT: Kumar, Anant Madan  
; TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE  
; FILE REFERENCE: SELL MH1116-PCT  
; CURRENT APPLICATION NUMBER: PCT/US01/26675  
; CURRENT FILING DATE: 2001-08-27  
; PRIOR APPLICATION NUMBER: 60/228,262  
; PRIOR FILING DATE: 2000-08-25  
; NUMBER OF SEQ ID NOS: 101  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
PCT-US01-26675-3

Query Match 100.0%; Score 666; DB 1; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 60  
|||||  
DB 39 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 98  
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 117  
|||||  
DB 99 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 155

## RESULT 4

US-09-119-209-2  
; Sequence 2, Application US/09119209  
; GENERAL INFORMATION:  
; APPLICANT: LASKY, LAURENCE A.  
; APPLICANT: STACHELL, SCOTT E.  
; APPLICANT: ROSEN, STEVEN D.  
; APPLICANT: SINGER, MARK S.  
; APPLICANT: YEDNOCK, TED A.  
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Winpatin (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,209  
; FILING DATE: 20-Jul-1998  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/513278  
; FILING DATE: 10-AUG-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/059027  
; FILING DATE: 6-MAY-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/786149  
; FILING DATE: 31-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/315015  
; FILING DATE: 23-FEB-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Love, Richard B.  
; REGISTRATION NUMBER: 34,659  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650/225-5530  
; TELEFAX: 650/952-9881  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 amino acids  
; TYPE: Amino Acid  
; TOPOLOGY: Linear  
US-09-119-209-2

Query Match 100.0%; Score 666; DB 15; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 60  
|||||  
DB 39 WTYHSEKPMNQRARFCRDNYTDLVAIONKAEIEYLEKTLPSRSYWIIGIRKIGIW 98  
QY 61 TWVGINKSLTEAEENWGDGEPNNKKNKEDCVEIYIKRNKDGAKWDDACHKRLKALC 117

Db 99 TWVGTNKSLLTEBAENWGDEPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLKALC 155

RESULT 5  
US-09-760-475-2123  
; Sequence 2123, Application US/09760475  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT249  
; CURRENT APPLICATION NUMBER: US/09/760,475  
; PRIORITY FILING DATE: 2001-01-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4122  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2123  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-760-475-2123

Query Match 100.0%; Score 666; DB 21; Length 382;  
Best Local Similarity 100.0%; Pred. No. 1,1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 60  
Db 49 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 108  
Oy 61 TWVGTNKSLLTEBAENWGDEPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLKALC 117  
Db 109 TWVGTNKSLLTEBAENWGDEPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLKALC 165

RESULT 6  
PCT-US92-03970-2  
; Sequence 2, Application PC/TUS9203970  
; GENERAL INFORMATION:  
; APPLICANT: Dana-Farber Cancer Institute, Inc.  
; TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: Massachusetts  
; COUNTRY: U.S.A.  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/03970  
; FILING DATE: 19920513  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Helne, Holliday C.  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: DFCI-152Bq9  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; TELEX: 940675  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear

; MOLECULE TYPE: protein  
PCT-US92-03970-2

Query Match 100.0%; Score 666; DB 1; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1,1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 60  
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 111  
Oy 61 TWVGTNKSLLTEBAENWGDEPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLKALC 117  
Db 112 TWVGTNKSLLTEBAENWGDEPNKKKEDCVETIYIKRNKDAGKWNDDACHKRLKALC 168

RESULT 7  
PCT-US94-00909-2  
; Sequence 2, Application PC/TUS9400909  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING  
; NUMBER OF SEQUENCES: 11  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (Epo)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US94/00909  
; PRIOR APPLICATION DATA:  
; FILING DATE: 25-JAN-1993  
; APPLICATION NUMBER: US 08/008,459  
; PRIOR APPLICATION DATA:  
; FILING DATE: 30-NOV-1992  
; APPLICATION NUMBER: US 07/983,606  
; PRIOR APPLICATION DATA:  
; FILING DATE: 03-OCT-1991  
; APPLICATION NUMBER: US 07/770,608  
; PRIOR APPLICATION DATA:  
; FILING DATE: 02-APR-1992  
; APPLICATION NUMBER: US 07/962,483  
; PRIOR APPLICATION DATA:  
; FILING DATE: 29-JUL-1991  
; APPLICATION NUMBER: US 07/737,092  
; PRIOR APPLICATION DATA:  
; FILING DATE: 08-JUL-1991  
; APPLICATION NUMBER: US 07/730,503  
; PRIOR APPLICATION DATA:  
; FILING DATE: 15-MAY-1991  
; APPLICATION NUMBER: US 07/700,773  
; PRIOR APPLICATION DATA:  
; FILING DATE: 21-FEB-1989  
; APPLICATION NUMBER: US 07/313,109  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US94-00909-2

Query Match 100.0%; Score 666; DB 1; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1,1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 60  
Db 52 WTYHSEKPMNQRRARFCRDNYTDLVAIONKAEIYEKTLPSFSRYWIGIRKIGTW 111

Qy 61 TWGTNKSSTLEAEENMGDEPNKKNKEDCVEIYIKRNKDGAKNDACHKLAALC 117  
|||||  
Db 112 TWGTNKSSTLEAEENMGDEPNKKNKEDCVEIYIKRNKDGAKNDACHKLAALC 168

RESULT 8  
US-08-008-459-2  
Sequence 2, Application US/08008459  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/008,459  
FILING DATE: 25-JAN-1993  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/983,606  
FILING DATE: 30-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/862,483  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,608  
FILING DATE: 03-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/737,092  
FILING DATE: 29-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/730,503  
FILING DATE: 08-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/700,773  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/313,109  
FILING DATE: 21-FEB-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Heine, Holliday C.  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: DFCI-318XX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
TELEX: 940675  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-008-459-2

Query Match 100.0%; Score 666; DB 4; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1,1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 WTYHYSEKPMNOMARRFCDNVTDLVAIONKAEIETLEKTLFPSRSYTWIGIRKIGIM 60

Db 52 WTYHYSEKPMNOMARRFCDNVTDLVAIONKAEIETLEKTLFPSRSYTWIGIRKIGIM 111  
|||||  
Qy 61 TWGTNKSSTLEAEENMGDEPNKKNKEDCVEIYIKRNKDGAKNDACHKLAALC 117  
|||||  
Db 112 TWGTNKSSTLEAEENMGDEPNKKNKEDCVEIYIKRNKDGAKNDACHKLAALC 168

RESULT 9  
US-08-340-539-2  
Sequence 2, Application US/08340539  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,539  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/983,606  
FILING DATE: 30-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/862,483  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,608  
FILING DATE: 03-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/737,092  
FILING DATE: 29-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/730,503  
FILING DATE: 08-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/700,773  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/313,109  
FILING DATE: 21-FEB-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Heine, Holliday C.  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: DFCI-318XX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
TELEX: 940675  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539-2

Query Match 100.0%; Score 666; DB 7; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 60  
|||||  
DB 52 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 111  
|||||  
QY 61 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 117  
|||||  
DB 112 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 168  
|||||

## RESULT 10

US-08-410-569-2  
; Sequence 2, Application US/08410569  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Sperlin, Olivier G.  
; TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/410,569  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,608  
; FILING DATE: 03-OCT-1991  
; APPLICATION NUMBER: US 07/700,773  
; FILING DATE: 15-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Helne, Holliday C.  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: DFCG-152EX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; TELEX: 940675  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-410-569-2

Query Match 100.0%; Score 666; DB 8; Length 385;  
Best Local Similarity 100.0%; Pred. No. 1.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 60  
|||||  
DB 52 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 111  
|||||  
QY 61 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 117  
|||||  
DB 112 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 168  
|||||

RESULT 11  
US-60-207-315-428  
; Sequence 428, Application US/60207315  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; TITLE OF INVENTION: USES THEREOF  
; FILE REFERENCE: C1000601  
; CURRENT APPLICATION NUMBER: US/60/207,315  
; CURRENT FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 528  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 428  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-207-315-428

Query Match 100.0%; Score 666; DB 26; Length 1078;  
Best Local Similarity 100.0%; Pred. No. 4.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 60  
|||||  
DB 621 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 680  
|||||  
QY 61 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 117  
|||||  
DB 681 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 737  
|||||

RESULT 12  
US-60-212-659-523  
; Sequence 523, Application US/60212659  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: C1000674  
; CURRENT APPLICATION NUMBER: US/60/212,659  
; CURRENT FILING DATE: 2000-06-19  
; NUMBER OF SEQ ID NOS: 879  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 523  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-212-659-523

Query Match 100.0%; Score 666; DB 26; Length 1078;  
Best Local Similarity 100.0%; Pred. No. 4.1e-67;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 60  
|||||  
DB 621 WTYHSEKPMNQARARFCRDNYTDLVAIQNKAIEYLEKTLPSRSYYWIGIRKIGIM 680  
|||||  
QY 61 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 117  
|||||  
DB 681 TWGTVNKSILTEAEENWGDEPNKKKEDCVETIYIKRNKDAGKWDACCHKIKALC 737  
|||||

RESULT 13  
US-60-230-435-1751  
; Sequence 1751, Application US/60230435  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: USES THEREOF

```

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; FILE REFERENCE: CL000768
; CURRENT APPLICATION NUMBER: US/60/230,435
; CURRENT FILING DATE: 2000-09-06
; NUMBER OF SEQ ID NOS: 2991
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1751
; LENGTH: 1078
; TYPE: PRT
; ORGANISM: HUMAN
US-60-230-435-1751
```

```

Query Match          100.0%; Score 666; DB 26; Length 1078;
Best Local Similarity 100.0%; Pred. No. 4.1e-67;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
DB 621 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 680
    |||||||
QY 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
DB 681 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 737
    |||||||
```

```

RESULT 14
US-60-160-189-8687
; Sequence 8687, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; FILE REFERENCE: CL000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8687
; LENGTH: 119
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-8687
```

```

Query Match          98.2%; Score 654; DB 26; Length 119;
Best Local Similarity 98.3%; Pred. No. 6e-67;
Matches 115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
QY 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
DB 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
```

```

RESULT 15
US-60-169-867-5823
; Sequence 5823, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: CL000160
```

```

; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5823
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-5823
```

```

Query Match          98.2%; Score 654; DB 26; Length 119;
Best Local Similarity 98.3%; Pred. No. 6e-67;
Matches 115; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
QY 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
DB 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
```

```

RESULT 16
US-60-160-203-5003
; Sequence 5003, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5003
; LENGTH: 119
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-5003
```

```

Query Match          95.0%; Score 633; DB 26; Length 119;
Best Local Similarity 94.9%; Pred. No. 1.6e-64;
Matches 111; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```

QY 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
DB 1 WTYHSEKPMNQARARFCRDNYTDVAIONKAEIYLEKTLPSRSYTWIGIRKIGIW 60
    |||||||
QY 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
DB 61 TWGINKSLTEAEENMGDEPNKKNKEDCVEIYIKRNKDAGKWNDDACHKLKALC 117
    |||||||
```

```

RESULT 17
US-60-169-840-6716
; Sequence 6716, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; FILE REFERENCE: CL000160
```



```
; FILE REFERENCE: CL000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 6716
; LENGTH: 119
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(119)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-6716

Query Match          95.0%; Score 633; DB 26; Length 119;
Best Local Similarity 94.9%; Pred. No. 1.6e-64;
Matches 111; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWGIRKIGIWM 60
    |||||||
DB 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWGIRKIGIWM 60
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDGKWNDACHKRLKALC 117
    |||||||
DB 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDGKWNDACHKRLKALC 117
    |||||||

RESULT 18
US-09-119-209-4
; Sequence 4, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209
; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
```

```
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-4

Query Match          84.8%; Score 565; DB 15; Length 372;
Best Local Similarity 82.9%; Pred. No. 4.8e-56;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWGIRKIGIWM 60
    |||||||
DB 39 WYHYSEKPMNMENARFKQNTDLVAIONKREIYLENTLPKSPYYWIGIRKIGKWM 98
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNKDGKWNDACHKRLKALC 117
    |||||||
DB 99 TWVGTNKTITKEAENMGAGEPNKKKEDCVELIYIKRERDSGKWNDACHKRLKALC 155
    |||||||

RESULT 19
US-60-160-189-10011
; Sequence 10011, Application US/60160189
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000112
; CURRENT APPLICATION NUMBER: US/60/160,189
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 10162
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10011
; LENGTH: 116
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-189-10011

Query Match          80.9%; Score 539; DB 26; Length 116;
Best Local Similarity 96.0%; Pred. No. 1.1e-53;
Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWGIRKIGIWM 60
    |||||||
DB 11 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIYLEKTLPEFSRSYWGIRKIGIWM 70
    |||||||

QY 61 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKRNK 99
    |||||||
DB 71 TWVGTNKSLEAEENMGDEPNKKKEDCVELIYIKEQ 109
    |||||||

RESULT 20
US-60-160-203-6200
; Sequence 6200, Application US/60160203
; GENERAL INFORMATION:
; APPLICANT: BONAZZI, VIVIEN
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000116
; CURRENT APPLICATION NUMBER: US/60/160,203
; CURRENT FILING DATE: 1999-10-19
; NUMBER OF SEQ ID NOS: 6374
; SOFTWARE: FastSeq for Windows Version 4.0
```

```
; SEQ ID NO 6200
; LENGTH: 116
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-160-203-6200
```

Query Match 80.9%; Score 539; DB 26; Length 116;  
Best Local Similarity 96.0%; Pred. No. 1.1e-53;

Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 70
    |||
OY 61 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRKQ 109
    |||
```

## RESULT 21

```
US-60-169-840-9326
; Sequence 9326, Application US/60169840
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000164
; CURRENT APPLICATION NUMBER: US/60/169,840
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 9628
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9326
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-840-9326
```

Query Match 80.9%; Score 539; DB 26; Length 116;  
Best Local Similarity 96.0%; Pred. No. 1.1e-53;

Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 70
    |||
OY 61 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRKQ 109
    |||
```

## RESULT 22

```
US-60-169-867-7998
; Sequence 7998, Application US/60169867
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS,
; FILE REFERENCE: C1000160
; CURRENT APPLICATION NUMBER: US/60/169,867
; CURRENT FILING DATE: 1999-12-09
; NUMBER OF SEQ ID NOS: 8230
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7998
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Human
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(116)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-169-867-7998
```

Query Match 80.9%; Score 539; DB 26; Length 116;  
Best Local Similarity 96.0%; Pred. No. 1.1e-53;

Matches 95; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    |||
DB 11 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 70
    |||
OY 61 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRK 99
    |||
DB 71 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRKQ 109
    |||
```

## RESULT 23

```
US-60-207-315-467
; Sequence 467, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: C1000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467
; LENGTH: 700
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(700)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-467
```

Query Match 66.8%; Score 445; DB 26; Length 700;  
Best Local Similarity 65.8%; Pred. No. 7.5e-42;

Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

```
OY 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAEIYLEKTLPFSSRYWIGIRKIGIW 60
    |||
DB 34 WTYHSTFAYSNISIRKCCQNKRYTDLVAIONKNEIDYINKYLPYSSRYWIGIRKKNKT 93
    |||
OY 61 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKRKADGKKNDDACHLKAALC 117
    |||
DB 94 TWGINKSLTEAEENMGDEPNKKKEDCVEIYIKSPADGKKNDDCHLKAALC 150
    |||
```

## RESULT 24

```
PCT-US94-09395-4
; Sequence 4, Application PC/TUS9409395
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
```

```

Query Match          66.8%; Score 445; DB 1; Length 830;
Best Local Similarity 65.8%; Pred. No. 9.3e-42;
Matches 77; Conservative 11; Mismatches 29; Indels 0; Gaps 0;

Oy      1 WTYHSEKPMNQARRCRDNYTDLVALIQNKAEIETLEKTLPFERSYYWIGIRKIGIWM 60
        ||||| : : : : : ||||| || : || : ||||| || : ||||| ||
Db      42 WTYHSTKASVSNIRKQCNRYTDLVALIQNKNEIDYLNKVIPLYSSYWIGIRKNNKTW 101
Oy      61 TWVGNKSLTEBAEWMGCGEPNNKKNKEDCVIITYIKRNKDAGKNYDDACHIKKALC 117
        ||||| : || ||||| ||||| ||||| ||||| ||||| : || ||||
Db      102 TWVGNKSLTEBAEWMADNEPNNKRNNEDEVETIKSPAGKMWDEHCKRKHALC 158

RESULT 25
US-08-449-6878-4
; Sequence 4, Application US/084496878
; GENERAL INFORMATION:
; APPLICANT: McEever, Rodger P.
; APPLICANT: Pan, Junliang
; TITLE OF INVENTION: Expression Control Sequences of the
; TITLE OF INVENTION: P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/449,687B
; FILING DATE: 24-MAY-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/110,158
; FILING DATE: 20-AUG-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/320,408
; FILING DATE: 08-MAR-1989

```

	Query Match	66.8%	Cred. 445;	DB 8;	Length 830;
	Best Local Similarity	65.8%	Pred. No. 9.3e-42;		
	Matches	77;	Conservative 11;	Mismatches 29;	Indels 0; Gaps 0;
Oy	1 WTHYSEKPMNQRRRCRDNYTPTLVAIQNAELEYLEKTLPFSRSYYWIGIRKIGTWI	60			
	. : : : :				
Db	42 WTHYSTAYSNISRIKKYCQNRITPTLVALQNNEIDYLNKTYPIVYSYYWIGIRNNKTW	101			
Oy	61 TWVGNTSLTEAEKNMGDGPENNNKKEDCEVTIYIKRNKDAGKNMDDACHIKALAC	117			
Db	102 TWVGSKRLTNLAENWADNPEPNKKRNNDECCVEITYIKSPAPAKWNDEHCKLKKAHALC	158			

```

Query Match          66.8%  Score 445;  DB 24;  Length 830;
Best Local Similarity 65.8%  Pred. No. 9, 3e-42;
Matches 77;  Conservative 11;  Mismatches 29;  Indels 0;  Gaps 0;

Qy      1  WTYHSEKPMNQARRRPFCDNYTDLVAIQNAETIELEKTPFSSRYWIGIRKIGGIW 60
        |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      42  WTYHSTKAYSWNISRKYCQNYRTDVAIQNNKEIDYLNKPKYPPSSYWIGIRKNNKTW 101

Qy      61  TWVGNKSLSTEEAEMWGQSEPPNNKKKEEDCVETIYIRNNKDAGWYNDACCHKKAALC 117
        |||||  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      102  TWVGNKSLSTEEAEMWADNEPNNKRNNEEDCVETIYIKSPAGKWNDEHCKLKHALLC 158

RESULT  27
US-60-196-718-4236
Sequence 4236, Application US/60196718
: GENERAL INFORMATION:
: APPLICANT: Bonazzzi, Vivien
: TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF

```

```
; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4236
; LENGTH: 129
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4236
```

## Query Match

Best Local Similarity 66.1%; Score 440.5; DB 26; Length 129;  
Matches 77; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

```
QY 1 WTYHSEKPMNOMORARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRYWYIGIRKIGTW 60
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 6 WTYHSTKAYSWE-SRKYCONRYTDLVAIONKNEIDLNVLPYSSYTWIGIRKNNKTW 65
QY 61 TWGVTNKSILTEAEWNGDGEPPNNKKKEDCVEIYIKRNKDGAKWMDACHKLKALC 117
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 66 TWGTRKALITNEAEWADNEPNNKRNEDCVEIYIKESV-SGKWNDEHCLKKHALC 121
```

## RESULT 28

```
US-60-196-718-4237
; Sequence 4237, Application US/60196718
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4237
; LENGTH: 133
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4237
```

## Query Match

Best Local Similarity 66.1%; Score 440.5; DB 26; Length 133;  
Matches 77; Conservative 13; Mismatches 26; Indels 1; Gaps 1;

```
QY 1 WTYHSEKPMNOMORARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRYWYIGIRKIGTW 60
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 17 WTYHSTKAYSWE-SRKYCONRYTDLVAIONKNEIDLNVLPYSSYTWIGIRKNNKTW 76
QY 61 TWGVTNKSILTEAEWNGDGEPPNNKKKEDCVEIYIKRNKDGAKWMDACHKLKALC 117
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 77 TWGTRKALITNEAEWADNEPNNKRNEDCVEIYIKESV-SGKWNDEHCLKKHALC 132
```

## RESULT 29

```
US-60-195-053-1909
; Sequence 1909, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1909
; LENGTH: 128
```

```
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-1909
```

## Query Match

Best Local Similarity 66.0%; Score 439.5; DB 26; Length 128;  
Matches 77; Conservative 12; Mismatches 27; Indels 1; Gaps 1;

```
QY 1 WTYHSEKPMNOMORARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRYWYIGIRKIGTW 60
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 11 WTYHSTKAYSWE-SRKYCONRYTDLVAIONKNEIDLNVLPYSSYTWIGIRKNNKTW 69
QY 61 TWGVTNKSILTEAEWNGDGEPPNNKKKEDCVEIYIKRNKDGAKWMDACHKLKALC 117
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 70 TWGTRKALITNEAEWADNEPNNKRNEDCVEIYIKESV-SGKWNDEHCLKKHALC 126
```

## RESULT 30

```
US-60-195-053-1908
; Sequence 1908, Application US/60195053
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000427
; CURRENT APPLICATION NUMBER: US/60/195,053
; CURRENT FILING DATE: 2000-04-06
; NUMBER OF SEQ ID NOS: 2836
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1908
; LENGTH: 129
; TYPE: PRT
; ORGANISM: HUMAN
US-60-195-053-1908
```

## Query Match

Best Local Similarity 66.0%; Score 439.5; DB 26; Length 129;  
Matches 77; Conservative 12; Mismatches 27; Indels 1; Gaps 1;

```
QY 1 WTYHSEKPMNOMORARFCRDNYTDLVAIONKAEIEYLEKTLPEFSRYWYIGIRKIGTW 60
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 6 WTYHSTKAYSWE-SRKYCONRYTDLVAIONKNEIDLNVLPYSSYTWIGIRKNNKTW 64
QY 61 TWGVTNKSILTEAEWNGDGEPPNNKKKEDCVEIYIKRNKDGAKWMDACHKLKALC 117
    ||||| | : : : : : ||||| | : : : : : ||||| | : : : : : ||||| | : : : : :
Db 65 TWGTRKALITNEAEWADNEPNNKRNEDCVEIYIKESV-SGKWNDEHCLKKHALC 121
```

## RESULT 31

```
US-60-196-718-4238
; Sequence 4238, Application US/60196718
; GENERAL INFORMATION:
; APPLICANT: Bonazzi, Vivien
; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND
; FILE REFERENCE: CL000456
; CURRENT APPLICATION NUMBER: US/60/196,718
; CURRENT FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 7494
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4238
; LENGTH: 138
; TYPE: PRT
; ORGANISM: HUMAN
US-60-196-718-4238
```

Query Match 66.0%; Score 439.5; DB 26; Length 138;  
Best Local Similarity 65.8%; Pred. No. 4.1e-42;







**This Page Blank (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:10:07 : Search time 28.63 Seconds  
(without alignments)  
99.818 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_39\_155

Perfect score: 666

Sequence: 1 WTYHYSEKPMNQBARRECR.....NKDAGKWDACHTKALIC 117

Scoring table:

BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Issued Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	666	100.0	117	1	US-08-274-661B-38 Sequence 38, Appl
2	666	100.0	372	2	US-08-513-278-2 Sequence 2, Appl
3	666	100.0	372	6	5514582-2 Patent No. 5514582
4	666	100.0	385	1	US-08-340-539A-2 Sequence 2, Appl
5	666	100.0	385	2	US-08-461-592B-2 Sequence 2, Appl
6	575	86.3	119	1	US-08-340-539A-14 Sequence 14, Appl
7	565	84.8	117	6	5514582-7 Patent No. 5514582
8	565	84.8	372	2	US-08-513-278-4 Sequence 4, Appl
9	565	84.8	372	6	5514582-4 Patent No. 5514582
10	560	84.1	119	1	US-08-340-539A-12 Sequence 12, Appl
11	557	83.6	119	1	US-08-340-539A-13 Sequence 13, Appl
12	455	66.3	119	1	US-08-340-539A-18 Sequence 18, Appl
13	447	67.1	830	5	PCR-0591-05059-2 Sequence 2, Appl
14	447	67.1	830	6	5378464-2 Patent No. 5378464
15	445	66.8	117	1	US-08-274-661B-39 Sequence 39, Appl
16	445	66.8	119	1	US-08-340-539A-16 Sequence 16, Appl
17	445	66.8	119	1	US-08-340-539A-17 Sequence 17, Appl
18	445	66.8	830	1	US-08-110-158-4 Sequence 4, Appl
19	444	66.7	120	1	US-08-340-539A-19 Sequence 19, Appl
20	434	65.2	119	1	US-08-274-661B-37 Sequence 37, Appl
21	425	63.8	484	2	US-08-252-493C-9 Sequence 9, Appl
22	425	63.8	484	3	US-09-276-197-9 Sequence 9, Appl
23	415	62.3	119	1	US-08-340-539A-15 Sequence 15, Appl
24	415	62.3	120	1	US-08-274-661B-36 Sequence 36, Appl
25	415	62.3	610	1	US-08-365-470-3 Sequence 3, Appl
26	415	62.3	610	3	US-09-209-668-19 Sequence 19, Appl
27	415	62.3	610	4	US-09-009-490A-89 Sequence 89, Appl

28	415	62.3	610	6	5217870-2	Patent No. 5217870
29	237	35.6	67	3	US-08-840-062-8	Sequence 8, Appl
30	164	24.6	38	2	US-08-513-278-5	Sequence 5, Appl
31	164	24.6	38	6	5514582-5	Patent No. 5514582
32	161	24.2	31	2	US-08-140-137A-44	Sequence 44, Appl
33	160.5	24.1	273	4	US-09-111-470-10	Sequence 10, Appl
34	160.5	24.1	292	2	US-08-688-342-4	Sequence 4, Appl
35	160.5	24.1	292	2	US-09-113-788-4	Sequence 4, Appl
36	160.5	24.1	316	4	US-09-111-470-4	Sequence 4, Appl
37	154.5	23.2	110	6	5514582-12	Patent No. 5514582
38	154.5	23.2	174	1	US-07-641-971B-1	Sequence 1, Appl
39	154.5	23.2	174	1	US-07-781-248A-1	Sequence 1, Appl
40	154.5	23.2	320	1	US-08-365-103B-10	Sequence 10, Appl
41	154.5	23.2	321	1	US-08-365-103B-8	Sequence 8, Appl
42	150.5	22.6	287	4	US-09-111-470-6	Sequence 6, Appl
43	149.5	22.4	114	6	5514582-8	Patent No. 5514582
44	149	22.4	111	6	5514582-11	Patent No. 5514582
45	148	22.2	1479	3	US-08-840-062-2	Sequence 2, Appl

#### ALIGNMENTS

```
RESULT 1
US-08-274-661B-38
; Sequence 38, Application US/08274661B
; Patent No. 5593882
; GENERAL INFORMATION:
; APPLICANT: Erbe, David V.
; APPLICANT: Lasky, Laurence A.
; APPLICANT: Presta, Leonard G.
; TITLE OF INVENTION: Selectin Variants
; NUMBER OF SEQUENCES: 39
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Winpatin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/274,661B
; FILING DATE: 13-Jul-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/956701
; FILING DATE: 10/01/1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
; REFERENCE/DOCKET NUMBER: 761P1C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-3216
; TELEFAX: 415/952-9881
; TELEX: 910/371-7168
; INFORMATION FOR SEQ ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 117 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-08-274-661B-38

Query Match 100.0%; Score 666; DB 1; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.8e-66;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 WTYHYSEKPMNQBARRECRDNTDVAIONKAEIYLEKTLPEFSRSYIWIGIRKIGIWM 60
```

Db 1 WTHYSEKPNMWRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60  
QY 61 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 117  
Db 61 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 117

RESULT 2  
US-08-513-278-2  
; Sequence 2, Application US/08513278  
; Patent No. 5840844  
; GENERAL INFORMATION:  
; APPLICANT: LASKY, LAURENCE A.  
; APPLICANT: STACHELL, SCOTT E.  
; APPLICANT: ROSEN, STEVEN D.  
; APPLICANT: SINGER, MARK S.  
; APPLICANT: YEDNOCK, TED A.  
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 460 Point San Bruno Blvd  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: patln (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/513,278  
; FILING DATE: 10-AUG-1993  
; CLASSIFICATION: 5530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/059027  
; FILING DATE: 06-MAY-1993  
; APPLICATION NUMBER: 07/786149  
; FILING DATE: 31-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/315015  
; FILING DATE: 23-FEB-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dreger, Ginger R.  
; REGISTRATION NUMBER: 33,055  
; REFERENCE/DOCKET NUMBER: 565D1C1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415/225-3216  
; TELEFAX: 415/952-9881  
; TELEX: 910/371-7168  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 372 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; US-08-513-278-2

Query Match 100.0%; Score 666; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 7.5e-66;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTHYSEKPNMWRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60  
Db 39 WTHYSEKPNMWRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 98  
QY 61 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 117  
Db 99 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 155

RESULT 3  
5514582-2  
; Patent No. 5514582  
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
; IMMNOGLOBULINS  
; NUMBER OF SEQUENCES: 43  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/185,670  
; FILING DATE: 21-JAN-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 986,931  
; FILING DATE: 08-DEC-1992  
; APPLICATION NUMBER: 808,122  
; FILING DATE: 16-DEC-1991  
; APPLICATION NUMBER: 440,625  
; FILING DATE: 22-NOV-1989  
; APPLICATION NUMBER: 315,015  
; FILING DATE: 23-FEB-1989  
; SEQ ID NO: 2:  
; LENGTH: 372  
5514582-2

Query Match 100.0%; Score 666; DB 6; Length 372;  
Best Local Similarity 100.0%; Pred. No. 7.5e-66;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTHYSEKPNMWRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 60  
Db 39 WTHYSEKPNMWRARFCNDNTDVAIONKAEIYLEKTLFSSSYWIGIRKIGIM 98  
QY 61 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 117  
Db 99 TWGJTKSLTEEAENMGDGEPPNNKKNEKDCVEIYIRKNDAGKWDACHKILKAALC 155

RESULT 4  
US-08-340-539A-2  
; Sequence 2, Application US/08340539A  
; Patent No. 5808025  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,539A  
; FILING DATE: 16-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,459  
; FILING DATE: 25-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gunnison, Jane  
; REGISTRATION NUMBER: 38,479  
; REFERENCE/DOCKET NUMBER: CG-104 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-2

Query Match 100.0%; Score 666; DB 1; Length 385;  
Best Local Similarity 100.0%; Pred. No. 7.8e-66;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 60  
|||||  
Db 52 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 111  
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 117  
|||||  
Db 112 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 168

RESULT 5  
US-08-461-592B-2  
Sequence 2, Application US/08461592B  
Patent No. 5834425

GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESS: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,592B  
FILING DATE:  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/340,539  
FILING DATE: 16-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: James F. Haley, Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: CG-104  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-461-592B-2

Query Match 100.0%; Score 666; DB 2; Length 385;  
Best Local Similarity 100.0%; Pred. No. 7.8e-66;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 60  
|||||  
Db 52 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 111  
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 117  
|||||  
Db 112 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 168

RESULT 6  
US-08-340-539A-14

Sequence 14, Application US/08340539A

Patent No. 5808025

GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESS: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9090  
TELEFAX: 212-596-9000  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-14

Query Match 86.3%; Score 575; DB 1; Length 119;  
Best Local Similarity 82.1%; Pred. No. 2.1e-56;  
Matches 96; Conservative 12; Mismatches 9; Indels 0; Gaps 0;

Qy 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 60  
|||||  
Db 1 WTYHSEKPMNQARARFCRDNYTDLVAIONKAETLEKTLPEFSRYWIGIRKIGTW 60  
Qy 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 117  
|||||  
Db 61 TWVGTNKSLEAEAWMGDEPNKKKEDCVELIYIKRNKDGKWNDDACHKLAALC 117

RESULT 7  
5514582-7  
Patent No. 5514582

```
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
IMMUNOGLOBULINS
NUMBER OF SEQUENCES: 43
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/185,670
FILING DATE: 21-JAN-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 986,931
FILING DATE: 08-DEC-1992
APPLICATION NUMBER: 808,122
FILING DATE: 16-DEC-1991
APPLICATION NUMBER: 440,625
FILING DATE: 22-NOV-1989
APPLICATION NUMBER: 315,015
FILING DATE: 23-FEB-1989
SEQ ID NO: 7
LENGTH: 117
5514582-7
```

```
Query Match      84.8%; Score 565; DB 6; Length 117;
Best Local Similarity 82.9%; Pred. No. 2.7e-55;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLFPSRSYYWIGIRKIGIW 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 TWGINKSLTEAEWNGGEPNNKKNKEDCVEIYIKRKNKDGKWNDDACHKRAALC 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 61 TWGINKTLTKEAENWGAGEPNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
RESULT 8
US-08-513-278-4
; Sequence 4, Application US/08513278
; Patent No. 5840844
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/513,278
; FILING DATE: 10-AUG-1995
; CLASSIFICATION: 5530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 06-MAY-1993
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Dreger, Ginger R.
; REGISTRATION NUMBER: 33,055
```

```
REFERENCE/DOCKET NUMBER: 565D1C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415/225-3216
TELEFAX: 415/952-9881
TELEX: 910/371-7168
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 372 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-513-278-4
```

```
Query Match      84.8%; Score 565; DB 2; Length 372;
Best Local Similarity 82.9%; Pred. No. 1.1e-54;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLFPSRSYYWIGIRKIGIW 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 39 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 98
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 TWGINKSLTEAEWNGGEPNNKKNKEDCVEIYIKRKNKDGKWNDDACHKRAALC 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 99 TWGINKTLTKEAENWGAGEPNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 155
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
RESULT 9
5514582-4
; Patent No. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 4
; LENGTH: 372
5514582-4
```

```
Query Match      84.8%; Score 565; DB 6; Length 372;
Best Local Similarity 82.9%; Pred. No. 1.1e-54;
Matches 97; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAEIETLEKTLFPSRSYYWIGIRKIGIW 60
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 39 WTYHSEKPMNMENARKFCOKNTYDLVAIONKREIETLENTLPKSPYYWIGIRKIGKW 98
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 61 TWGINKSLTEAEWNGGEPNNKKNKEDCVEIYIKRKNKDGKWNDDACHKRAALC 117
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 99 TWGINKTLTKEAENWGAGEPNKKSKEDCVEIYIKRERDSGKWNDDACHKRAALC 155
    |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

```
RESULT 10
US-08-340-539A-12
; Sequence 12 Application US/08340539A
; Patent No. 5608025
; GENERAL INFORMATION:
; APPLICANT: Tedder, Thomas F.
; APPLICANT: Kansas, Geoffrey S.
; TITLE OF INVENTION: CHIMERIC SELECTING AS SIMULTANEOUS
; BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
; NUMBER OF SEQUENCES: 28
```



LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-18

Query Match 68.3%; Score 455; DB 1; Length 119;  
Best Local Similarity 67.5%; Pred. No. 3.9e-43;  
Matches 79; Conservative 10; Mismatches 28; Indels 0; Gaps 0;

QY 1 WTHYSEKPNMORARPCRDNTDVAIONKAEIYLEKTLPEFSRYSYIGIRKIGIWM 60  
DB 1 WTHYSTKAVSMNSRVCPCRRHPTDVAIONKNEIAHLNDVIFENSYWIGIRKINMKW 60  
QY 61 TWVGTKSLTEAEANGDEPNKKKNEKEDCEVEIYIKRNNDAGWMDACHIKALC 117  
DB 61 TWVGTKSLTEAEANGDEPNKKKNEKEDCEVEIYIKRNNDAGWMDACHIKALC 117

RESULT 13  
PCT-US91-05059-2

Sequence 2, Application PC/TUS9105059

GENERAL INFORMATION:

APPLICANT: Regents of the Board of the, University of

APPLICANT: Oklahoma

TITLE OF INVENTION: Functionally Active Selectin-Derived

TITLE OF INVENTION: Peptides

NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick & Cody

STREET: 100 Peachtree Street, Suite 3100

CITY: Atlanta

STATE: Georgia

COUNTRY: US

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US91/05059

FILING DATE: 19910717

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/320408

FILING DATE: 08-MAR-1989

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 07/554199

FILING DATE: 17-JUL-1990

ATTORNEY/AGENT INFORMATION:

NAME: Pabst, Patricia L.

REGISTRATION NUMBER: 31,284

REFERENCE/DOCKET NUMBER: OMRP110CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: 404-572-6508

TELEFAX: 404-572-6555

INFORMATION FOR SEQ. ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 830 amino acids

TYPE: AMINO ACID

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: YES

ANTI-SENSE: YES

FRAGMENT TYPE: N-terminal

ORGANISM: Homo sapien

TISSUE TYPE: Blood

CELL TYPE: Endothelial

FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 4..25  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 60..158  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 131..150  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 163..174  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 168..183  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 185..194  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 200..244  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 213..226  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 230..257  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 262..306  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 275..288  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 292..319  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 324..368  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 337..350  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 354..381  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 386..430  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 399..412  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 416..443  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 448..492  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 461..474  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 478..505  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 510..554  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 523..536  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 540..567



? APPLICATION NUMBER: 07/956701  
?  
? FILING DATE: 10/01/1992  
?  
? ATTORNEY/AGENT INFORMATION:  
? NAME: Dieger, Ginger R.  
? REGISTRATION NUMBER: 33,055  
? REFERENCE/DOCKET NUMBER: 761PICG  
? TELECOMMUNICATION INFORMATION:  
? TELEPHONE: 415/225-3216  
? TELEFAX: 415/952-9881  
? TELERX: 910/371-7168  
?  
? INFORMATION FOR SEQ ID NO: 39:  
?  
? SEQUENCE CHARACTERISTICS:  
? LENGTH: 117 amino acids  
? TYPE: Amino Acid  
? TOPOLOGY: Linear

US-08-274-661B-39

Query Match	66.8%;	Score 445;	DB 1;	Length 117;
Best Local Similarity	65.8%;	Pred. No. 4.9e-42;		
Matches 77;	Conservative 11;	Mismatches 29;	Indels 0;	Gaps 0

[illegible]

RESULT 16  
US-08-340-539A-16  
; Sequence 16, Application US/08340539A  
; Patent No. 5000000

APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:

STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA

```

: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/340,539A  
 FILING DATE: 16-NOV-1994  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:

ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
FILING DATE: 25-JAN-1993  
APPLICATION NUMBER: US 08/008,453

REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 16:

LENGTH: 119 amino acids  
TYPE: amino acid  
STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: protein  
US-08-340-539A-16

Query Match	66.8%;	Score 445;	DB 1;	Length 119;
Best Local Similarity	65.8%;	Pred. No. 5e-42;		
Matches 77;	Conservative 14;	Mismatches 26;	Indels 0;	Gaps 0;

[illegible]

RESULT 17  
US-08-340-539A-17  
; Sequence 17, Application US/08340539A  
Patent No. 5808025

GENERAL INFORMATION: Thomas F.  
APPLICANT: Tedder,  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE

;;  
;;  
;;  
SIREET: 1251 Avenue of the Americas  
;;  
CITY: New York  
;;  
STATE: New York  
;;  
COUNTRY: USA  
;;  
ZIP: 10020

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER.: US/08/340,539A
ATTENTION:

```

CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:

NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SOURCE: 17

SEQUENCE CHARACTERISTICS:  
LENGTH: 119 amino acids  
TYPE: amino acid  
STANDARDNESS:

TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 MS-08-340-539A-17

Query Match	66.8%;	Score 445;	DB 1;	Length 119;
Best Local Similarity	66.0%;	Score 440;	DB 1;	Length 119;

Matches	77, Conservative	11; Mismatches	29; Indels	0; Gaps
1	WTFTYHSERDKMNTGAPADDECCPDNWTEDTATATGATKATATTTT	TTTAT	TTTAT	TTTAT

1 WTHTSKAYSMNISRKQCNRYTDLVAIQNKNEIDYLNKVLPPYSSYYWIGIRKNNK 60

61 TWGTSKSLTEAENWGDEPNKKNKEDCEIYIKRNKDAGKWNDDACHKLKALC 117





PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956701  
FILING DATE: 10/01/1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 120 amino acids  
TYPE: Amino acid  
TOPOLOGY: Linear  
US-08-274-661B-37

Query Match 65.2%; Score 434; DB 1; Length 120;  
Best Local Similarity 63.2%; Pred. No. 8,3e-41;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTHYSEKPMWQARRCRNDITDVAIONKAEIEYLEKTLPPSRSYTWIGIRKIGTW 60  
DB 1 WTHYSEKPMWQARRCRNDITDVAIONKAEIEYLEKTLPPSRSYTWIGIRKIGTW 60  
OY 61 TWGTSKSLTEAEENMGDEPNKKNEDEVEIYIKRNKDGKWNDDACHKKAALC 117  
DB 61 IAVGTHKPLTEGAKNMVAGEPNKNNQNNEDCEIYIKRKDGMNDRCSKKLALC 117

RESULT 21  
US-08-252-493C-9  
Sequence 9, Application US/08252493C  
Patent No. 5891645  
GENERAL INFORMATION:  
APPLICANT: Rollins, Scott  
APPLICANT: Rother, Russell P.  
APPLICANT: Evans, Mark J.  
APPLICANT: Matlis, Louis A.  
TITLE OF INVENTION: PORCINE E-SELECTIN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seth A. Fidel  
STREET: 25 Science Park, Box 15  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06511  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 750 Kb storage  
COMPUTER: PC compatible  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252,493C  
FILING DATE: June 1, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fidel, Seth A.  
REGISTRATION NUMBER: 38,449  
REFERENCE/DOCKET NUMBER: ALX-138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 776-1790  
TELEFAX: (203) 772-3655  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid

STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: Predicted amino acid sequence of  
DESCRIPTION: Porcine E-selectin  
US-08-252-493C-9

Query Match 63.8%; Score 425; DB 2; Length 484;  
Best Local Similarity 63.2%; Pred. No. 4,5e-39;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;

OY 1 WTHYSEKPMWQARRCRNDITDVAIONKAEIEYLEKTLPPSRSYTWIGIRKIGTW 60  
DB 23 WSYASTETWTFEDASAYCOORYTHLVAIONHAEIEYLNSTFVNSASYTWIGIRKIGTW 82  
OY 61 TWGTSKSLTEAEENMGDEPNKKNEDEVEIYIKRNKDGKWNDDACHKKAALC 117  
DB 83 TWIGTKALTPENATNMVAGEPNKKNSNEDCEIYIKRKDSGKWNDRCSKKLALC 139

RESULT 22  
US-09-276-197-9  
Sequence 9, Application US/09276197  
Patent No. 6040428  
GENERAL INFORMATION:  
APPLICANT: Rollins, Scott  
APPLICANT: Rother, Russell P.  
APPLICANT: Evans, Mark J.  
APPLICANT: Matlis, Louis A.  
TITLE OF INVENTION: PORCINE E-SELECTIN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seth A. Fidel  
STREET: 25 Science Park, Box 15  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06511  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 750 Kb storage  
COMPUTER: PC compatible  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/276,197  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/252,493  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Fidel, Seth A.  
REGISTRATION NUMBER: 38,449  
REFERENCE/DOCKET NUMBER: ALX-138  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (203) 776-1790  
TELEFAX: (203) 772-3655  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 484 amino acids  
TYPE: amino acid  
STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: cDNA to mRNA  
DESCRIPTION: Predicted amino acid sequence of  
DESCRIPTION: Porcine E-selectin  
US-09-276-197-9

Query Match 63.8%; Score 425; DB 3; Length 484;  
Best Local Similarity 63.2%; Pred. No. 4,5e-39;  
Matches 74; Conservative 13; Mismatches 30; Indels 0; Gaps 0;







APPLICANT: CARON, DANIEL J.; LASKY, LAURENCE A.  
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 43  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,670  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 986,931  
FILING DATE: 08-DEC-1992  
APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989  
SEQ ID NO: 5:  
LENGTH: 38  
5514582-5

Query Match 24.6%; Score 164; DB 6; Length 38;  
Best Local Similarity 78.4%; Pred. No. 1.4e-11;  
Matches 29; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 2 TYHSEKPMNQARFCRDNTDVAIONKAEIETLPS 38  
DB 2 TYHSEKPMNENARKFKYONTDLVAIONKXXIETL 38

RESULT 32  
US-08-140-137A-44  
Sequence 44, Application US/08140137A  
Patent No. 5817617  
GENERAL INFORMATION:  
APPLICANT: TUOMANEN, ELAINE  
APPLICANT: MASURE, H. R.  
TITLE OF INVENTION: ANALOGS OF ENDOTHELIAL LEUKOCYTE  
ADHESION MOLECULE (ELAM)  
NUMBER OF SEQUENCES: 49  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Klauber & Jackson  
STREET: 411 Hackensack Avenue  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/140,137A  
FILING DATE: 27-MAY-1994  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-096  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201 487-5800  
TELEFAX: 201 343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
DESCRIPTION: hLHR peptide sequence  
US-08-140-137A-44

Query Match 24.2%; Score 161; DB 2; Length 31;  
Best Local Similarity 100.0%; Pred. No. 2.4e-11;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 15 ARRCRDNTDVAIONKAEIETLPS 45  
DB 1 ARRCRDNTDVAIONKAEIETLPS 31

RESULT 33  
US-09-111-470-10  
Sequence 10, Application US/09111470  
Patent No. 6277959  
GENERAL INFORMATION:  
APPLICANT: Valladeau, Jenny  
APPLICANT: Ravel, Odile  
APPLICANT: Bates, Elizabeth E.M.  
APPLICANT: Ford, John  
APPLICANT: Saeland, Sem  
APPLICANT: Lebecque, Serge J.E.  
TITLE OF INVENTION: Mammalian Membrane Protein Genes;  
TITLE OF INVENTION: Related Reagents  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DNAX Research Institute  
STREET: 901 California Avenue  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304-1104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/111,470  
FILING DATE: 08-JUL-1998  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/053,080  
FILING DATE: 09-JUL-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Ching, Edwin P.  
REGISTRATION NUMBER: 34,090  
REFERENCE/DOCKET NUMBER: SF0695  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)852-9196  
TELEFAX: (650)496-1200  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 273 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-111-470-10

Query Match 24.1%; Score 160.5; DB 4; Length 273;  
Best Local Similarity 29.6%; Pred. No. 3.9e-10;  
Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

OY 3 YHSEKPMNQARFCRDNTDVAIONKAEIETLPSFSYWIIGRTIGITW 62  
DB 150 YWFSHSGSMWAIEKYOLKNAHLVINSREONFVOKYL--GSAIYWGSLSPDEGAKW 207  
OY 63 V-GTKSLTEEAENWGDGEPNNKRNK-----EDCVELIYIRNKDAGKWDACGK 111  
DB 208 VDTGTD--YATGFQWKKPGQPDMDQGHGIGGEDCAHFH-----PDGRMNDVQGR 255

RESULT 34



```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/111,470
; FILING DATE: 08-JUL-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/053,080
; FILING DATE: 09-JUL-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: CHING, Edwin P.
; REGISTRATION NUMBER: 34,090
; REFERENCE/DOCKET NUMBER: SF0695
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650)852-9196
; TELEFAX: (650)496-1200
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 316 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-111-470-4
```

```

Query Match          24.1%; Score 160.5; DB 4; Length 316;
Best Local Similarity 29.6%; Pred. No. 4.6e-10;
Matches 34; Conservative 22; Mismatches 44; Indels 15; Gaps 5;

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWIIGIRKIGITW 62
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 193 YFSGHSMGMAEAEKYCQKKNHLYVINSREONVQKYL--GSAYTWGGLSDPGAKW 250
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 63 V-GTKSLTEAEENWGDEPNKKKKK-----EDCVEIYIKRNKDGAKWDDACRK 111
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 251 VDGTD--YATGFQNNKPGQPDWQCHGLGGECAHFH-----PDGRWNNDDVCCR 298
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 37
5514582-12
; PATENT NO. 5514582
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
; IMMUNOGLOBULINS
; NUMBER OF SEQUENCES: 43
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,670
; FILING DATE: 21-JAN-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 986,931
; FILING DATE: 08-DEC-1992
; APPLICATION NUMBER: 808,122
; FILING DATE: 16-DEC-1991
; APPLICATION NUMBER: 440,625
; FILING DATE: 22-NOV-1989
; APPLICATION NUMBER: 315,015
; FILING DATE: 23-FEB-1989
; SEQ ID NO: 12:
; LENGTH: 110
; 5514582-12
```

```

Query Match          23.2%; Score 154.5; DB 6; Length 110;
Best Local Similarity 32.2%; Pred. No. 5.8e-10;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWIIGIRKIG--GIW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 3 YYFGKGTQWVHARACDDMEGLVSIHSPEDFLTKHSHSGS--WIGLRNLDLKEEF 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 TWVGTNKSLEAEENWGDEPNKKKKKEDCVEIYIKRNKDGAKWDDAC--HKLKALC 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 61 IWDGSH--VDISNNAPEPTSRSGEDCVAM-----RSGRWNDACFDRKLGAMVC 110
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 38
US-07-641-971B-1
; Sequence 1, Application US/07/641971B
; Patent No. 5236706
; GENERAL INFORMATION:
; APPLICANT: Debire, Patrice
; APPLICANT: Mossalayi, Mohammed D
; TITLE OF INVENTION: A PHARMACEUTICAL PREPARATION FOR THE
; MATURETION OF PROTHYMOCYTES
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
; STREET: 556 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/641,971B
; FILING DATE: 19910116
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 90016254
; FILING DATE: 24-JAN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Fishman, Irving M
; REGISTRATION NUMBER: 30258
; REFERENCE/DOCKET NUMBER: 4-17921/+/DEB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-277-4832
; TELEFAX: 908-277-4306
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; ORIGINAL SOURCE:
; CELL TYPE: Human B. Cells
; CELL LINE: CHO cells transformed with pCAL8-BF-ND
; US-07-641-971B-1
```

```

Query Match          23.2%; Score 154.5; DB 1; Length 174;
Best Local Similarity 32.2%; Pred. No. 1e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMNORARRCRDNYTDVAIONKAEIEYLEKTLPSRSYWIIGIRKIG--GIW 60
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 28 YYFGKGTQWVHARACDDMEGLVSIHSPEDFLTKHSHSGS--WIGLRNLDLKEEF 85
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 61 TWVGTNKSLEAEENWGDEPNKKKKKEDCVEIYIKRNKDGAKWDDAC--HKLKALC 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 86 IWDGSH--VDISNNAPEPTSRSGEDCVAM-----RSGRWNDACFDRKLGAMVC 135
   | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```

RESULT 39
US-07-781-248A-1
; Sequence 1, Application US/07/781248A
; Patent No. 5246599
; GENERAL INFORMATION:
; APPLICANT: Debire, Patrice
; APPLICANT: Mossalayi, Mohammed D
; TITLE OF INVENTION: MATURATION OF HEMATOPOIETIC CELLS
; NUMBER OF SEQUENCES: 6
```



```

CORRESPONDENCE ADDRESS:
ADDRESSEE: Irving M. Fishman, CIBA-GEIGY Corporation
STREET: 556 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/81,248A
FILING DATE: 19911230
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 90103565
FILING DATE: 09-MAY-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ikeler, Barbara J.
REGISTRATION NUMBER: 36,170
REFERENCE/DOCKET NUMBER: 4-18065/A/DEB
TELECOMMUNICATION INFORMATION:
TELEPHONE: 908-277-3368
TELEFAX: 908-277-4306
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: peptide
HYPOTHEICAL: NO
ANTI-SENSE: NO
FRAGMENT TYPE: N-terminal
ORIGINAL SOURCE:
CELL TYPE: Human B. Cells
CELL LINE: CHO cells transformed with pCAL8-BF-ND
US-07-781-248A-1

Query Match      23.2%; Score 154.5; DB 1; Length 174;
Best Local Similarity 32.2%; Pred. No. 1e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMWQRRRCRDNYDVAIQKAEIYLEKTLPPRSYYWIGIRKIG--GITW 60
   1::: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 28 YFEGKTKQWVHARYACDMEGOLVSHSPEDDFLTKHASHTGS--WIGLRNLDKGEF 85
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 61 TWVGTKSLTEAEKNGDSEPNKKKKEDCVETIYIRKNDAGKWNDDAC-HKTKAALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 86 IWDGSH---VDYSNAPGEPTSRSGEDCVMM-----RSGRWNDAPCDRLGAWVC 135
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

RESULT 40
US-08-365-103B-10
Sequence 10 Application US/08365103B
Patent No. 576943
GENERAL INFORMATION:
APPLICANT: Lynch, Richard G.
APPLICANT: Nunez, Raphael D.
APPLICANT: Yodot, Jungi
TITLE OF INVENTION: DNA Sequences for Soluble Forms of CD23
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Zarley, McKee, Thombe, Voorhees & Sease
STREET: 801 Grand Ave. Suite 3200
CITY: Des Moines
STATE: Iowa
COUNTRY: United States
ZIP: 50309
COMPUTER READABLE FORM:

```

```

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,103B
FILING DATE: 28-DEC-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Nebel, Heidi S.
REGISTRATION NUMBER: 37,719
REFERENCE/DOCKET NUMBER: Ulf N5-24
TELECOMMUNICATION INFORMATION:
TELEPHONE: (515) 288-3667
TELEFAX: (515) 288-1338
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 320 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-365-103B-10

Query Match      23.2%; Score 154.5; DB 1; Length 320;
Best Local Similarity 32.2%; Pred. No. 2.2e-09;
Matches 38; Conservative 20; Mismatches 47; Indels 13; Gaps 5;

QY 3 YHSEKPMWQRRRCRDNYDVAIQKAEIYLEKTLPPRSYYWIGIRKIG--GITW 60
   1::: 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 174 YFEGKTKQWVHARYACDMEGOLVSHSPEDDFLTKHASHTGS--WIGLRNLDKGEF 231
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

QY 61 TWVGTKSLTEAEKNGDSEPNKKKKEDCVETIYIRKNDAGKWNDDAC-HKTKAALC 117
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
Db 232 IWDGSH---VDYSNAPGEPTSRSGEDCVMM-----RSGRWNDAPCDRLGAWVC 281
   1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1

```

Search completed: September 7, 2002, 10:14:42  
Job time: 275 sec

**This Page Blank (uspto)**



FT	Modified-site	269..271
FT	/label= glycosylation_site	
FT	Modified-site	285..287
FT	/label= glycosylation_site	
FT	Modified-site	299..301
FT	/label= glycosylation_site	
FT	Modified-site	324..326
FT	/label= glycosylation_site	
XX	US\$506126-A.	
PN		
PD	09-APR-1996.	
XX		
PF	25-FEB-1988;	88US-0160416.
XX		
PR	01-DEC-1992;	92US-0983647.
PR	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	13-JUL-1990;	90US-0553759.
PR	18-OCT-1993;	93US-0139273.
XX		
PA	(GEHO ) GEN HOSPITAL CORP.	
XX		
PI	Arufo A, Seed B;	
XX		
DR	WPI; 1996-200279/20.	
DR	N-PsDB; AA#14723.	
XX		
PT	Cloning of cDNA encoding cell surface antigen - useful for isolation of diagnostic and therapeutic proteins	
PS		
PX	Example 14; Column 71-74; 79pp; English.	
XX		
CC	2 cDNA clones encoding Leu8 determinants (AA#91442 and AA#91443) were isolated from a human T-cell library using a novel immunoselection cloning method. The longer insert (AA#14723) contained 2,350 residues, while the shorter lacked 436 internal residues. A major transcript of 2.4 kb was present in peripheral blood mononuclear cells (PBMC), tonsillar B cells, and several lymphocytic cell lines, and a minor transcript of 2.0 kb in PBMC, Jurkat and HSb-2 leukaemic T-cell lines. Leu8 antigens can be obtd. for diagnostic and therapeutic use. The presence or absence of Leu8 on CD4+ T-cells identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets. Soluble forms of Leu8 can act as antiinflammatory agents by reducing lymphocyte migration.	
CC		
CC		
CC		
SQ	Sequence	363 AA:
Query Match	100.0%;	Score 666; DB 17; Length 363;
Best Local Similarity	100.0%;	Pred. No. 2.2e-63;
Matches 117; Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
OY	1 WTYHSEKMMNQRARPCRDNYTDLVALIONKAETLEYEKTPEFSRYWGIRKIGGW 60       52 wtyhsekpmmqgrarrfcdydtlvaltqmnaaeylektpfsrywgiirigtw 111	
OY	61 TWVGTKNSLTFEEAENMGDEEPNNKNKKEDCVELIYIKRNKGAKWNDDACHKLKAALC 117       Db 112 twvgtnsklteeaenmgdgpnnknkedcveilyikrnkgakwnddachklkaalc 168	
RESULT 3		
ID	AA#96184 standard; Protein; 363 AA.	
XX	AA#96184;	
XX	AA#96184;	
DT	19-DEC-2000 (first entry)	
XX	Human T-cell specific Leu8 antigen.	
XX	Leu8; cell surface antigen; human; immunoselection; panning;	

KW Immunodiagnosis; diagnosis; immunotherapy; gene therapy;  
 KW Immune disorder; infection; aschma; immune-complex disease;  
 KW amyloidosis; multiple sclerosis; inflammation; antinflammatory.

OS Homo sapiens.

PN US6111093-A.

XX 29-AUG-2000.

PF 28-OCT-1996; 980US-0181612.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 23-MAR-1990; 90US-0498809.

PR 13-JUL-1990; 90US-0553759.

XX (GEHO ) GEN HOSPITAL CORP.

PI Stamenkovic I, Seed B;

XX WPI: 2000-586382/55.

DR N-PSDB: AAA50632.

PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,  
 PT useful for immunodiagnosis and immunotherapy of immune-mediated  
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic  
 PT diseases

XX Example 14; Column 69-72; 75pp: English.

CC The present sequence is that of a T-cell specific Leu8 antigen, as  
 CC predicted from 1 of 2 clones (see AAA50632) isolated from a human  
 CC T-cell library by the method of the invention. A longer Leu8  
 CC antigen is given in AA96138. The method, designed to isolate cell  
 CC surface antigen (CSA) cDNAs, is based upon transient expression of  
 CC a CSA in eukaryotic cells and physical selection of cells expressing  
 CC the antigen by adhesion to (panning on) an antibody-coated substrate  
 CC such as a culture dish. CSA nucleic acids isolated by the method of  
 CC the invention, and the proteins they encode, are useful for  
 CC immunodiagnostic and immunotherapeutic applications, including the  
 CC diagnosis and treatment of immune-mediated infections, diseases, and  
 CC disorders in animals, including humans. These disorders include  
 CC asthma, immune-complex disease, amyloidosis, parasitic diseases or  
 CC multiple sclerosis. The ability to interfere with the binding of  
 CC Leu8-T-cells to antigen presenting cells, or the ability to cause  
 CC such binding to occur on surfaces other than lymphocyte cells, can  
 CC be useful in diagnostics and therapy. The level of activated Leu8-  
 CC T-cells relative to resting Leu8+ cells could serve as a measure of  
 CC immune response to a particular antigen. Modification of the  
 CC specificity of the extracellular domain of Leu8, which mediates  
 CC adhesion to specific endothelial cells of lymph nodes, could serve  
 CC to regulate the homing potential of resting T cells. Soluble forms  
 CC of Leu8 could act as antiinflammatory agents by reducing lymphocyte  
 CC migration.

XX Sequence 363 AA:

Query Match 100.0%; Score 666; DB 21: Length 363;  
 Best Local Similarity 100.0%; Pred. No. 2.2e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTHYSEKPMNQARRCRDNYTDVAIONKAETLEKTLPFSSRYWIGIRKIGTM 60  
 DB 52 WYHYSEKPMNQARRCRDNYTDVAIONKAETLEKTLPFSSRYWIGIRKIGTM 111  
 QY 61 TWVGINKSLTEFAENKGDSEPPNNKKKCEVEIYIKRNKDAGKWDNDACHKIKAALC 117  
 DB 112 TWVGINKSLTEFAENKGDSEPPNNKKKCEVEIYIKRNKDAGKWDNDACHKIKAALC 168

RESULT 4

AA98127  
 ID AAR98127 standard; Protein; 370 AA.

XX AAR98127;

DT 01-NOV-1996 (first entry)

XX Human lymphocyte cell surface glycoprotein (HULHR) variant.

XX Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

XX metastasis.

OS Homo sapiens.

XX Key

XX Region

XX Domain

XX Domain

XX Binding-site

XX Binding-site

XX Domain

XX Domain

XX US5514582-A.

XX 07-MAY-1996.

XX 23-FEB-1989; 89US-0315015.

XX 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX (GEHTH ) GENENTECH INC.

PA Capon DJ, Lasky LA;

XX WPI: 1996-238773/24.

XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand

XX binding site of a receptor fused to Ig constant region - useful for

XX diagnosis and treatment e.g. of inflammation

XX Disclosure; Page 19; 41pp; English.

XX A hybrid immunoglobulin chain comprising the ligand binding site of

XX a single transmembrane receptor without an active transmembrane

XX region; fused at its C-terminus with the N-terminus of an

XX immunoglobulin constant region. The receptor is not a member of the

XX immunoglobulin super family, nor a multiple subunit polypeptide

XX encoded by discrete genes. The hybrid immunoglobulin chain combines

XX the adhesion/targeting of a ligand binding partner (LBP) with the

XX effector functions of immunoglobulin and can bind to and/or activate

XX more than one ligand. It can be used diagnostically for the in

XX vitro assay of LBP and their targets; or therapeutically to deliver

XX LBP such as toxins, enzymes, growth factors to particular cells.

XX Typical applications are as antiviral, neuromodulating and

XX immunomodulating agents, or as modulators of cell adhesion (e.g. in

XX treating graft rejection; inflammation; metastasis of lymphoma etc.)

XX The immunoglobulin component increases plasma half life and

XX facilitates purification while deletion of the transmembrane region

PT binding site of a receptor fused to Ig constant region - useful for  
 PR diagnosis and treatment e.g. of inflammation  
 PS  
 XX Disclosure; Page 19; 41pp; English.  
 CC A hybrid immunoglobulin chain...

Disclosure; Page 19; 41pp; English.

CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region, fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. *Ab* antibody

the hybrid immunoglobulin chain combines the heavy chain of the antibody with the *adhesion/targeting* of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the *in vitro* assay of LBP and their targets, or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection, inflammation, metastasis of epithelial

CC The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein, CD45, are described in this report.

cc  
near surface glycoprotein described in AAR9106 are given in  
AAR9109-R9135. This variant is a Ser220 deletion mutant.  
XX  
Sequence 370 AA;  
50

Query Match 100.0%: Score 666: DB 17: Month 370.

0y 1 WTYHYSEKPMNQARRRCRDNTDVAIQNKAEIEYLEKTLPFSRSYYWIGIRKGITW 60  
best locally 100.0%; read no. 2,3e-03;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0.

Db 38 wtyhysekpmwqrarrfcrdnytdlvaigknaeieylektlpfsrsywigirklgilw

61 TWGTTNKSLEAEWNGDEPNKKKKEDCEIYIKRNKDAGKWNDDACHKLKALC 117

98 lwvglnksl leadenwgdgepnknknkedcvely lkrnkdgkwnddachklkaalc 154

## RESULT 6

ID	AAR98130 standard; Protein; 370 AA.
----	-------------------------------------

AC AAR98130;  
vv

DT 01-NOV-1996 (first entry)  
XX

Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
transmembrane receptor; adhesion; targeting;

KW metastasis.

OS Homo sapiens.

EH	Key	Location/Qualifiers
ET	Region	1 37

	/Label=	Signal region.
F1	38-154	
F1T		
Domain		

FT	Domain	/label=seccli uoalln. 159..192
----	--------	-----------------------------------

FT	Binding-site	196..257

FT	258.315
Binding-site	

Domain	331..353
E1T	
E2T	

PI	DOI:10.1002/	354 . . 3/0
----	--------------	-------------

FT /Label= Cytoplasmic domain.  
XX US514582-A.  
PN 07-MAY-1996.  
XX  
XX 23-FEB-1989; 89US-0315015.  
XX  
XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GETH ) GENENTECH INC.  
XX Capon DJ, Lasky LA;  
XX WPI; 1996-238773/24.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
PS Disclosure; Page 19; 41pp; English.  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant is an Ile296 deletion mutant.  
XX  
XX Sequence 370 AA;  
SQ  
Query Match 100.0%; Score 666; DB 17; Length 370;  
Best Local Similarity 100.0%; Pred. No. 2,3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 WTVHYSEKPMQMORARRCBDNYTDLVAIONKAEIEYEKTPFSRSYYWIGIRKIGTW 60  
Db 38 WcYhSeKpMwqRrtrrtrrtdytlvaIqnaeeylektpIstSYWYgItrIggIw 97  
OY 61 TWVGTKNSLGTGEAENMGDEPNKKKDKCEVETIKRNKDAGKMNDDCHKLKALC 117  
Db 98 twvgtnkslteaenmgdgpnkknkdcveIyIkrnkdgkwnddachkklkaalc 154  
RESULT 7  
AAR98114 ID AAR98114 standard; Protein; 371 AA.  
XX  
XX AAR98114;  
XX 01-NOV-1996 (first entry)  
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Region 1..37  
FT /Label= Signal region.  
FT Domain 38..154  
FT /Label= Lectin domain.  
FT Domain 159..192  
FT /Label= EGF domain.  
FT Binding-site 196..257  
FT /Label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /Label= Complement binding repeat 2.  
FT Domain 332..354  
FT /Label= Transmembrane domain.  
FT Domain 355..371  
FT /Label= Cytoplasmic domain.  
XX  
XX US514582-A.  
PN 07-MAY-1996.  
XX  
XX 23-FEB-1989; 89US-0315015.  
XX  
XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GETH ) GENENTECH INC.  
XX Capon DJ, Lasky LA;  
XX WPI; 1996-238773/24.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
PS Disclosure; Page 19; 41pp; English.  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains an His168Gln substitution.  
XX  
XX Sequence 371 AA;  
SQ  
Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2,3e-63;





PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPI: 1996-238773/24.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX  
PS Disclosure: Page 19; 41pp; English.  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains an Asn181Gln substitution.  
XX  
SQ Sequence 371 AA;  
  
Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAIELEYLEKTPFSRSYWIIGIKIGIW 60  
Db 38 WYHYSEKPMNQRRRCRDNYTDLVAIONKAIELEYLEKTPFSRSYWIIGIKIGIW 97  
Oy 61 TWVGTKNSLTTEAEKMGDEPNKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 117  
Db 98 TWVGTKNSLTTEAEKMGDEPNKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 154  
  
RESULT 10  
AAR98117  
ID AAR98117 standard; Protein: 371 AA.  
XX  
AC AAR98117;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1..37  
FT Region /Label= Signal region.

FT Domain 38..154  
FT /Label= Lectin domain.  
FT Domain 159..192  
FT /Label= EGF domain.  
FT Binding-site 196..257  
FT /Label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /Label= Complement binding repeat 2.  
FT Domain 332..354  
FT /Label= Transmembrane domain.  
FT Domain 355..371  
FT /Label= Cytoplasmic domain.  
PN US514582-A.  
PD 07-MAY-1996.  
XX  
PF 23-FEB-1989; 89US-0315015.  
XX  
PR 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPI: 1996-238773/24.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX  
PS Disclosure: Page 19; 41pp; English.  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a Thr211Ser substitution.  
XX  
SQ Sequence 371 AA;  
  
Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Oy 1 WTYHSEKPMNQRRRCRDNYTDLVAIONKAIELEYLEKTPFSRSYWIIGIKIGIW 60  
Db 38 WYHYSEKPMNQRRRCRDNYTDLVAIONKAIELEYLEKTPFSRSYWIIGIKIGIW 97  
Oy 61 TWVGTKNSLTTEAEKMGDEPNKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 117  
Db 98 TWVGTKNSLTTEAEKMGDEPNKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 154

RESULT 11  
AAR98118  
ID AAR98118 standard; Protein; 371 AA.  
XX  
AC AAR98118;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FH Region 1..37  
FT /label= Signal region.  
FT 38..154  
FT /label= Lectin domain.  
FT 159..192  
FT /label= EGF domain.  
FT Binding-site 196..257  
FT /label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /label= Complement binding repeat 2.  
FT 332..354  
FT /label= Transmembrane domain.  
FT 355..371  
FT /label= Cytoplasmic domain.  
XX  
PN US5514582-A.  
XX  
PD 07-MAY-1996.  
XX  
PE 23-FEB-1989; 89US-0315015.  
XX  
PR 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPI; 1996-238773/24.  
XX  
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX  
PS Disclosure; Page 19; 41pp; English.  
XX  
CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family; nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a Phe214Leu substitution.  
XX  
SQ Sequence 371 AA;  
XX  
Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WTYHYSEKPMWGORARRECRDNTDVAIQNKAIEYLEKTPSPSYTWIGIRKIGGIW 60  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
38 WcYnysekpmwqarrifrcdnyldvaigkkaeylektpfsrsywigirkiqgw 97  
QY 61 TWVGTNKSITFEAEENMGDGEPPNNKKNKEDCEVEIYIKRNKDAGKWNDDACHIKKAALC 117  
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
98 twvgtnkslteeaenwgdgeppnnkknkdeceveiyikrnkdagkwnddachikkaalc 154  
RESULT 12  
AAR98119  
ID AAR98119 standard; Protein; 371 AA.  
XX  
AC AAR98119;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FH Region 1..37  
FT /label= Signal region.  
FT 38..154  
FT /label= Lectin domain.  
FT Domain 159..192  
FT /label= EGF domain.  
FT Binding-site 196..257  
FT /label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /label= Complement binding repeat 2.  
FT 332..354  
FT /label= Transmembrane domain.  
FT 355..371  
FT /label= Cytoplasmic domain.  
XX  
PN US5514582-A.  
XX  
PD 07-MAY-1996.  
XX  
PE 23-FEB-1989; 89US-0315015.  
XX  
PR 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPI; 1996-238773/24.  
XX

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation  
 XX  
 PS Disclosure: Page 19, 41pp; English.  
 XX  
 CC A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets, or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains a Ser226Thr substitution.  
 XX  
 SQ Sequence 371 AA;  
 Query Match 100.0%; Score 666; DB 17; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 WYHYSEKPMNQRRRCRDNYDVLAIQNAELEYEKLTPFSRSYWGIRKIGTW 60  
 Db 38 WYHYSEKPMNQRRRCRDNYDVLAIQNAELEYEKLTPFSRSYWGIRKIGTW 97  
 Qy 61 TWVGNNKSLTEFEANWGDGEPNNKKNKEDCVETIKRNKDGAKMDDCHIKALC 117  
 Db 98 TWVGNNKSLTEFEANWGDGEPNNKKNKEDCVETIKRNKDGAKMDDCHIKALC 154  
 RESULT 13  
 AAR98120  
 ID AAR98120 standard; Protein: 371 AA.  
 AC AAR98120;  
 XX  
 DT 01-NOV-1996 (first entry)  
 XX  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT 1..37 /label= Signal region.  
 FT 38..154 /label= Lectin domain.  
 FT /label= Lectin domain.  
 FT 159..192 /label= EGF domain.  
 FT 196..257 /label= Complement binding repeat 1.  
 FT 258..316 /label= Complement binding repeat 2.  
 FT 332..354 /label= Transmembrane domain.  
 FT Domain

FT Domain 355..371  
 FT /label= Cytoplasmic domain.  
 XX  
 FM US5514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.  
 XX  
 PA (GENE) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI, 1996-238773/24.  
 XX  
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation  
 XX  
 PS Disclosure: Page 19, 41pp; English.  
 XX  
 CC A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets, or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains a Phe244Met substitution.  
 XX  
 SQ Sequence 371 AA;  
 Query Match 100.0%; Score 666; DB 17; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 WYHYSEKPMNQRRRCRDNYDVLAIQNAELEYEKLTPFSRSYWGIRKIGTW 60  
 Db 38 WYHYSEKPMNQRRRCRDNYDVLAIQNAELEYEKLTPFSRSYWGIRKIGTW 97  
 Qy 61 TWVGNNKSLTEFEANWGDGEPNNKKNKEDCVETIKRNKDGAKMDDCHIKALC 117  
 Db 98 TWVGNNKSLTEFEANWGDGEPNNKKNKEDCVETIKRNKDGAKMDDCHIKALC 154  
 RESULT 14  
 AAR98121  
 ID AAR98121 standard; Protein: 371 AA.  
 AC AAR98121;  
 XX  
 DT 01-NOV-1996 (first entry)  
 XX  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

XX	Immunoglobulin; transmembrane receptor; adhesion; targeting; diagnosis; therapy; drug delivery; antiviral; neuromodulator; immunomodulator; cell adhesion; graft rejection; inflammation;
KW	metastasis.
KW	Homo sapiens.
XX	
OS	
XX	
XX	Key
FT	Region
FT	1..37
FT	/label= Signal region.
FT	38..154
FT	/label= Lectin domain.
FT	159..192
FT	/label= BGF domain.
FT	196..257
FT	/label= Complement binding repeat 1.
FT	258..316
FT	/label= Complement binding repeat 2.
FT	332..354
FT	/label= Transmembrane domain.
FT	355..371
FT	/label= Cytoplasmic domain.
XX	
XX	Domain
XX	
XX	US5514582-A.
XX	
XX	07-MAY-1996.
XX	
XX	23-FEB-1989; 89US-0315015.
XX	
XX	23-NOV-1989; 89US-0440625.
XX	23-FEB-1989; 89US-0315015.
XX	16-DEC-1991; 91US-0808121.
XX	08-DEC-1992; 92US-0986931.
XX	21-JAN-1994; 94US-0185670.
XX	
XX	(GETH ) GENENTECH INC.
XX	
XX	Capon DJ, Lasky LA;
XX	
XX	WP1; 1996-238773/24.
XX	
XX	Nucleic acid encoding hybrid immunoglobulin comprising the ligand binding site of a receptor fused to Ig constant region - useful for diagnosis and treatment e.g. of inflammation
XX	
XX	Disclosure; Page 19; 41pp; English.
XX	
XX	A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region; fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the in vitro assay of LBP and their targets, or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particularly to deliver typical applications are as antiviral, neuromodulating and neuromodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX	The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR98106 are given in AAR98109-R98135. This variant contains a Thr282Ser substitution.
XX	
XX	Sequence 371 AA:

Query Match 100.0%; Score 666; DB 17; Length 371;

Best Local Similarity	100.0%;	Pred. No. 2.3e-63;			
Matches 117;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
Qy	1	WTHTSEKPMWQARRFCRDNTDVAIONKAEIEYLEKTLPEFSRSYYWIGIRIGTGM	60		
Db	38	wlhysekpmwqarrfcrdnytdlvaigkaeieylektlpfsrsywigirigtw	97		
Qy	61	TWVGTNKSLEFEANMGDGEPPNKKNKEDCVELYIKRNNDAGKWNDACHKLKAALC	117		
Db	98	twvgtlnksleeeenmgdgeppnknknedcvelyikrnxdagkwnddachklkaalc	154		
RESULT 15					
AAR98122					
ID	AAR98122	standard; Protein; 371 AA.			
XX					
AC	AAR98122;				
XX					
DT	01-NOV-1996	(first entry)			
XX					
DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.				
XX					
KW	Immunoglobulin; transmembrane receptor; adhesion; targetting;				
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;				
KW	immunomodulator; cell adhesion; graft rejection; inflammation;				
XX	metastasis.				
OS					
XX	Homo sapiens.				
FH					
FT	Key	Location/Qualifiers			
FT	Region	1..37			
		/label= Signal region.			
FT	Domain	38..154			
		/label= Lectin domain.			
FT	Domain	159..192			
		/label= EGF domain.			
FT	Binding-site	196..257			
		/label= Complement binding repeat 1.			
FT	Binding-site	258..316			
		/label= Complement binding repeat 2.			
FT	Domain	332..354			
		/label= Transmembrane domain.			
FT	Domain	355..371			
		/label= Cytoplasmic domain.			
XX					
XX	US5514582-A.				
XX					
XX	07-MAY-1996.				
PD					
XX					
PF	23-FEB-1989;	89US-0315015.			
XX					
PR	22-NOV-1989;	89US-0440625.			
	23-FEB-1989;	89US-0315015.			
PR	16-DEC-1991;	91US-0808122.			
PR	08-DEC-1992;	92US-0986931.			
PR	21-JAN-1994;	94US-0185670.			
XX					
PA	(GETH ) GENENTECH INC.				
XX					
PI	Capon DJ, Lasky LA;				
XX					
DR	WPI; 1996-238773/24.				
XX					
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand				
PT	binding site of a receptor fused to Ig constant region - useful for				
PT	diagnosis and treatment e.g. of inflammation				
XX					
PS	Disclosure; Page 19; 41pp; English.				
XX					
CC	A hybrid immunoglobulin chain comprising the ligand binding site of				
CC	a single transmembrane receptor without an active transmembrane				
CC	region; fused at its C-terminus with the N-terminus of an				
CC	immunoglobulin constant region. The receptor is not a member of the				

CC Immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-988135. This variant contains an Ile288Val substitution.  
XX  
SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVHYSEKPMNQRRARFCRDNYTDLVAIONKAEIEYELEKTLPFSSRYWIGIRKIGIM 60  
DB 38 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIEYELEKTLPFSSRYWIGIRKIGIM 97  
QY 61 TWVGTNKSITTEBAENWGDGEPNNKKNEKDCVEIYIRKNDAGKWNDDACHKIKALC 117  
DB 98 TWVGTNKSITTEBAENWGDGEPNNKKNEKDCVEIYIRKNDAGKWNDDACHKIKALC 154

RESULT 16  
AAR98123  
ID AAR98123 standard; Protein: 371 AA.  
XX  
AC AAR98123;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
OS Homo sapiens.  
XX  
XX

FT Key Location/Qualifiers  
FT Region 1..37  
FT /Label= Signal region.  
FT Domain 38..154  
FT /Label= Lectin domain.  
FT Domain 159..192  
FT /Label= EGF domain.  
FT Binding-site 196..257  
FT /Label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /Label= Complement binding repeat 2.  
FT Domain 332..354  
FT /Label= Transmembrane domain.  
FT Domain 355..371  
FT /Label= Cytoplasmic domain.  
XX  
XX  
XX US5514582-A.  
XX PN 07-MAY-1996.  
XX PD 23-FEB-1989; 89US-0315015.  
XX PF 22-NOV-1989; 89US-0440625.  
XX PR

PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GENTH ) GENENTECH INC.  
XX  
XX Capon DJ, Lasky LA;  
XX WPL; 1996-238773/24.  
XX  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
XX binding site of a receptor fused to Ig constant region - useful for  
XX diagnosis and treatment e.g. of inflammation  
XX  
XX Disclosure: Page 19; 41pp; English.  
XX  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
XX a single transmembrane receptor without an active transmembrane  
XX region; fused at its C-terminus with the N-terminus of an  
XX immunoglobulin constant region. The receptor is not a member of the  
XX immunoglobulin super family, nor a multiple subunit polypeptide  
XX encoded by discrete genes. The hybrid immunoglobulin chain combines  
XX the adhesion/targeting of a ligand binding partner (LBP) with the  
XX effector functions of immunoglobulin and can bind to and/or activate  
XX more than one ligand. It can be used diagnostically for the in  
XX vitro assay of LBP and their targets; or therapeutically to deliver  
XX LBP such as toxins, enzymes, growth factors to particular cells.  
XX Typical applications are as antiviral, neuromodulating and  
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in  
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)  
XX The immunoglobulin component increases plasma half life and  
XX facilitates purification while deletion of the transmembrane region  
XX facilitates recovery, improves aqueous solubility and removes  
XX potentially immunogenic epitopes. Variants of the human lymphocyte  
XX cell surface glycoprotein described in AAR98106 are given in  
XX AAR98109-988135. This variant contains a Lys298-Lys299; Arg-Arg  
XX substitution.  
XX  
SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTVHYSEKPMNQRRARFCRDNYTDLVAIONKAEIEYELEKTLPFSSRYWIGIRKIGIM 60  
DB 38 WYHYSEKPMNQRRARFCRDNYTDLVAIONKAEIEYELEKTLPFSSRYWIGIRKIGIM 97  
QY 61 TWVGTNKSITTEBAENWGDGEPNNKKNEKDCVEIYIRKNDAGKWNDDACHKIKALC 117  
DB 98 TWVGTNKSITTEBAENWGDGEPNNKKNEKDCVEIYIRKNDAGKWNDDACHKIKALC 154

RESULT 17  
AAR98124  
ID AAR98124 standard; Protein: 371 AA.  
XX  
AC AAR98124;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX  
XX  
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
XX immunomodulator; cell adhesion; graft rejection; inflammation;  
XX metastasis.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
XX  
XX  
XX

Db	98	twgtnkstleaeenwgddgpenknknedcveiyikrnkdsgkwndackhkaalc	154
		RESULT 18	
		AAR98129	
XX	ID	AAR98129 standard; Protein; 371 AA.	
XX	AC	AAR98129;	
XX	DT	01-NOV-1996 (first entry)	
XX	XX	Human lymphocyte cell surface glycoprotein (HLHR) variant.	
DE	XX	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
XX	XX	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KW	XX	immunomodulator; cell adhesion; graft rejection; inflammation;	
KM	XX	metastasis.	
OS	XX	Homo sapiens.	
XX	XX		
FH	Key	Location/Qualifiers	
FT	Region	1..37	
FT	FT	/label= Signal region.	
FT	Domain	38..154	
FT	Domain	/label= Lectin domain.	
FT	FT	159..192	
FT	FT	/label= EGF domain.	
FT	Binding-site	196..257	
FT	FT	/label= Complement binding repeat 1.	
FT	FT	258..315	
FT	Domain	/label= Complement binding repeat 2.	
FT	Domain	331..353	
FT	Domain	/label= Transmembrane domain.	
XX	XX	354..370	
XX	XX	/label= Cytoplasmic domain.	
XX	XX		
PN	US5514582-A.		
XX	XX		
PD	07-MAY-1996.		
XX	XX		
PF	23-FEB-1989;	89US-0315015.	
XX	XX		
PR	23-NOV-1989;	89US-0440635.	
PR	23-FEB-1989;	89US-0315015.	
PR	16-DEC-1991;	91US-0808122.	
PR	08-DEC-1992;	92US-0986931.	
PR	21-JAN-1994;	94US-0185670.	
PA	(GETH ) GENENTECH INC.		
XX	XX		
PI	Capon DJ, Lasky LA;		
XX	XX		
DR	WPI: 1996-238773/24.		
XX	XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand		
PT	binding site of a receptor fused to Ig constant region - useful for		
PT	diagnosis and treatment e.g. of inflammation		
XX	XX		
PS	Disclosure; Page 19; 41pp; English.		
XX	XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of		
CC	a single transmembrane receptor without an active transmembrane		
CC	region; fused at its C-terminus with the N-terminus of an		
CC	immunoglobulin constant region. The receptor is not a member of the		
CC	immunoglobulin super family, nor a multiple subunit polypeptide		
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines		
CC	the adhesion/targeting of a ligand binding partner (LBP) with the		
CC	effector functions of immunoglobulin and can bind to and/or activate		
CC	more than one ligand. It can be used diagnostically for the in		
CC	vitro assay of LBP and their targets; or therapeutically to deliver		
CC	LBP such as toxins, enzymes, growth factors to particularly cells.		
CC	Typical applications are as antiviral, neuromodulating and		
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in		

CC creating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant is a Asn271 deletion mutant.  
 CC  
 XX

SQ Sequence 371 AA;

Query Match 100.0%; Score 666; DB 17; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNQRRARFCRDNYTDLVATONKAETFEYLEKTLPEFSRXYWIGIRKIGTW 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 39 WCYHYSEKPMNQRRARFCRDNYTDLVATONKAETFEYLEKTLPEFSRXYWIGIRKIGTW 98  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 61 TWVGTKNSLTBEAENMGDEPNKKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 99 TWVGTKNSLTBEAENMGDEPNKKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 155

RESULT 19  
 AAR37960  
 ID AAR37960 standard; Protein; 372 AA.  
 XX  
 AC AAR37960;  
 DT 08-OCT-1993 (first entry)  
 XX  
 DE Human Lymphocyte Homing Receptor.  
 XX  
 KM HuLHR; lymphocyte binding inhibition; lymphoma metastasis;  
 KM transplant rejection; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide  
 FT /label= signal\_sequence  
 FT 39..372  
 FT /note= "Trp39 is probable N-terminus of mature LHR"  
 FT Domain  
 FT 39..155  
 FT /label= Lectin\_domain  
 FT 60..62  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site  
 FT 104..106  
 FT /note= "potential N-glycosylation site"  
 FT Domain  
 FT 160..193  
 FT /label= EGF\_domain  
 FT 177..179  
 FT /note= "potential N-glycosylation site"  
 FT Region  
 FT 197..258  
 FT /label= Complement\_Binding\_Repeat\_1  
 FT 216..218  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site  
 FT 232..234  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site  
 FT 246..248  
 FT /note= "potential N-glycosylation site"  
 FT Region  
 FT 259..317  
 FT /label= Complement\_Binding\_Repeat\_2  
 FT 271..273  
 FT /note= "potential N-glycosylation site"  
 FT Modified-site  
 FT 311..313  
 FT /note= "potential N-glycosylation site"  
 FT Domain  
 FT 333..335  
 FT /label= Transmembrane\_Domain  
 FT /note= "stop transfer sequence"  
 FT 356..372  
 FT /label= Cytoplasmic\_Domain

XX US5216131-A.  
 PN 01-JUN-1993.  
 XX  
 PD 23-FEB-1989; 89US-0315015.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 XX  
 PR 23-FEB-1989; 89US-0315015.  
 PR 31-OCT-1991; 91US-0786149.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX

XX Lasky LA, Rosen SD, Singer MS, Stachel SE, Yednock TA;  
 PI WPI: 1993-188588/23.  
 XX N-PSDB: AAQ43154.  
 DR

XX Human and murine lymphocyte homing receptors to treat graft  
 PT rejection and inflammation - comprise carbohydrate binding,  
 PT epidermal growth factor and complement binding domains  
 XX  
 PS Claim 1: Fig 1 and Fig 3; 32pp; English.

XX A human peripheral blood lymphocyte cDNA library in lambda gt10 was  
 CC screened with a 2.2kb EcoRI insert of the murine Mel14 antigen clone  
 CC (i.e. a murine LHR sequence). The largest EcoRI insert (2.2kb) was  
 CC isolated and sequenced. The ORF codes for 372 amino acids with a mol.  
 CC wt. of approximately 42,200. Comparison of the HuLHR amino acid  
 CC sequence with the murine LHR sequence (AAR37961) showed a high degree  
 CC of amino acid conservation in each of the LHR domains, e.g. 96% in  
 CC the transmembrane domain and 83% in the carbohydrate binding domain.  
 CC The LHRs could be used to compete with the normal binding of  
 CC lymphocytes to lymphoid tissue to treat inflammation or graft  
 CC rejection. They could also be used to control lymphoma metastasis  
 CC and to treat conditions involving lymphocyte accumulation.

SQ Sequence 372 AA;

Query Match 100.0%; Score 666; DB 14; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHYSEKPMNQRRARFCRDNYTDLVATONKAETFEYLEKTLPEFSRXYWIGIRKIGTW 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 39 WCYHYSEKPMNQRRARFCRDNYTDLVATONKAETFEYLEKTLPEFSRXYWIGIRKIGTW 98  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 OY 61 TWVGTKNSLTBEAENMGDEPNKKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 DB 99 TWVGTKNSLTBEAENMGDEPNKKKKEDCEVEIYIKRNKDAGKWDNDACHKIKALC 155

RESULT 20

AAR38908  
 ID AAR38908 standard; Protein; 372 AA.

XX AAR38908;

XX 11-JAN-1994 (first entry)

XX HuLHR.

XX Human; murine; lymphocyte; cell surface glycoprotein; homing receptor;

KM LHR; endothelium; lymphoid tissue; signal; domain; complement binding;  
 KM carbohydrate binding; epidermal growth factor-like; egf; intracellular;  
 KM transmembrane binding; cytoplasmic; ligand binding partner protein;  
 KM TMD; LBPP.

XX Homo sapiens.

OS

XX Key

FT Peptide

Location/Qualifiers  
 20..32  
 /note= "Signal peptide"

Query Match	Best Local Similarity	100.0%	Score 666	DR 14	Length 372
Matches 117	Conservative	0	Mismatches 0	Indels 0	Gaps 0
1	WTYHYSEKPMNQARRRRCRDNYTDLVALAIONAAETLEYEKTTPFSKSYWIGIRKIGTM	60			
Db	39 WYHYSEKPMNQARRRRCRDNYTDLVALAIONAAETLEYEKTTPFSKSYWIGIRKIGTM	98			
QY	61 TWVGVNKSITTEAEKMWGDESEPNKKNKKECEVETIYIKRNKDAGKWDNDACHIKAKALC	117			
Db	99 TWVGVNKSITTEAEKMWGDESEPNKKNKKECEVETIYIKRNKDAGKWDNDACHIKAKALC	155			
RESULT 21					
1D	AAAR76506				
XX	AAAR76506 standard; Protein: 372 AA.				
MC	AAAR76506;				
XX					

DT	01-DEC-1995	(first entry)
XX		
DE	Human LHR.	
XX		
KW	Lymphocyte homing receptor; lymphocyte cell surface glycoprotein;	
XX	LHR; ligand binding partner; immunoglobulin; constant region;	
KW	antibody engineering; immunomodulator.	
XX		
OS	Homo sapiens.	
XX		
PH	Key	Location/Qualifiers
FT	Peptide	20..32
FT		/label= Sig.peptide
FT		/note= "hydrophobic domain, may act as signal
FT		for insertion into the endoplasmic
FT		reticulum lumen"
FT	Protein	39..372
FT		label= Mat_protein
FT	Domain	39..334
FT		/note= "putative extracellular domain"
FT	Modified-site	60..62
FT		/label= N-glycosylation-site
FT	Modified-site	104..106
FT		/label= N-glycosylation-site
FT	Modified-site	177..179
FT		/label= N-glycosylation-site
FT	Modified-site	216..218
FT		/label= N-glycosylation-site
FT	Modified-site	232..234
FT		/label= N-glycosylation-site
FT	Modified-site	246..248
FT		/label= N-glycosylation-site
FT	Modified-site	271..273
FT		/label= N-glycosylation-site
FT	Modified-site	311..313
FT		/label= N-glycosylation-site
FT	Domain	335..357
FT		/note= "putative stop transfer or membrane anchor
FT		domain"
FT	Region	358..372
FT		/note= "putative intracellular region"
XX		
PN	US5428130-A.	
XX		
PD	27-JUN-1995.	
XX		
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI; 1995-240086/31.	
DR	N-PSDB; AA092802.	
XX		
PT	New hybrid ligand binding partner molecules - fused to immunoglobulin	
PT	constant region sequences to increase stability and in vivo plasma	
PT	half-life	
XX		
PS	Disclosure; Flg.1a-1c; 40pp; English.	
XX		
CC	A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10	
CC	cDNA library derived from human peripheral blood lymphocyte mRNA	
CC	obt. from primary cells. A cDNA clone encoding LHR was isolated.	
XX		
XQ	Sequence 372 AA;	



Query Match 100.0%; Score 666; DB 16; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLTPFSRSYWIIGIRKIGIW 60  
 |||  
 Db 39 WYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLTPFSRSYWIIGIRKIGIW 98  
 |||

OY 61 TWVGTNKSITTEAEKNMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 117  
 |||  
 Db 99 twvgtnkslteeaenwgdepnknkknkedcveilylkrndagkwnddachkikaalc 155

RESULT 22  
 AAR83050  
 ID AAR83050 standard; Protein; 372 AA.  
 AC AAR83050;  
 XX  
 DT 31-JAN-1996 (first entry)  
 XX  
 DE Human LHR`  
 XX  
 XX Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;  
 KM Immunoglobulin; Igg; constant region; receptor-mediated disease;  
 KW vector; plasma-life.  
 XX  
 OS Homo sapiens.

XX Location/Qualifiers  
 FH Key 20..32  
 FT Peptide /label= Sig\_peptide  
 FT 39..155  
 FT /label= Carbohydrate\_binding\_domain  
 FT 160..193  
 FT /label= Epidermal\_growth\_factor\_domain  
 FT 197..317  
 FT /label= Complement\_factor\_binding\_domain  
 FT 333..355  
 FT /label= Transmembrane\_binding\_domain  
 FT 356..372  
 FT Domain /label= Cytoplasmic\_domain  
 FT /note= "potential stop transfer sequence"

FT Modified-site 60..62  
 FT /label= N-glycosylation\_site  
 FT 104..106  
 FT Modified-site /label= N-glycosylation\_site  
 FT 177..179  
 FT Modified-site /label= N-glycosylation\_site  
 FT 216..218  
 FT Modified-site /label= N-glycosylation\_site  
 FT 232..234  
 FT Modified-site /label= N-glycosylation\_site  
 FT 248..248  
 FT Modified-site /label= N-glycosylation\_site  
 FT 271..273  
 FT Modified-site /label= N-glycosylation\_site  
 FT 311..313  
 FT Modified-site /label= N-glycosylation\_site

XX US5455165-A.  
 XX  
 PD 03-OCT-1995.  
 XX  
 XX  
 PD 23-FEB-1989; 89US-0315015.  
 XX  
 XX  
 PD 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185669.  
 XX  
 PA (GETH ) GENENTECH INC.

XX Capon DJ, Laaky LA;  
 PI  
 XX  
 DR WPI: 1995-350776/45.  
 DR N-PSDB: AAT05869.  
 PT Expression vector encoding fusion protein to increase plasma life -  
 PT compresses receptor ligand binding site and Ig constant region, for  
 PT treatment of receptor mediated disease  
 PS Disclosure: Fig 1; 42pp; English.  
 XX  
 CC A mouse LHR (Lymphocyte cell surface glycoprotein) cDNA clone was used  
 CC to screen an oligo-dT primed lambda gt10 cDNA library derived from  
 CC human peripheral blood lymphocyte mRNA obtd. from primary cells. A  
 CC 2.2 kb clone (sequence given in AAT05869) was isolated that encoded the  
 CC human LHR protein (AAR83050). LHR-IgG hybrids were constructed for use  
 CC in the targeting of therapeutic moieties to lymphoid tissue.  
 XX  
 SQ Sequence 372 AA:

Query Match 100.0%; Score 666; DB 16; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLTPFSRSYWIIGIRKIGIW 60  
 |||  
 Db 39 WYHYSEKPMNQRRRCRDNTDLVAIQNKAETLEYLTPFSRSYWIIGIRKIGIW 98  
 |||

OY 61 TWVGTNKSITTEAEKNMGDEPNKKKKEPCVEIYIKRNKDAGKWNDDACHKIKAAIC 117  
 |||  
 Db 99 twvgtnkslteeaenwgdepnknkknkedcveilylkrndagkwnddachkikaalc 155

RESULT 23  
 AAR98106  
 ID AAR98106 standard; Protein; 372 AA.  
 AC AAR98106;  
 XX  
 DT 31-OCT-1996 (first entry)  
 XX  
 DE Human Lymphocyte cell surface glycoprotein (HULHR).  
 XX  
 XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KM diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 XX  
 OS Homo sapiens.

XX Location/Qualifiers  
 FH Key 1..38  
 FT Region /label= Signal region.  
 FT 39..155  
 FT /label= Lectin domain.  
 FT 160..193  
 FT Domain /label= EGF domain.  
 FT 197..258  
 FT Binding-site /label= Complement binding repeat 1.  
 FT 259..317  
 FT Binding-site /label= Complement binding repeat 2.  
 FT 333..355  
 FT Domain /label= Transmembrane domain.  
 FT 356..372  
 FT Domain /label= Cytoplasmic domain.

XX US5514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 XX  
 PD 23-FEB-1989; 89US-0315015.  
 XX

XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
PA (GETH ) GENENTECH INC.  
XX  
XX Capon DJ, Lasky LA;  
XX  
XX WPI: 1996-238773/24.  
DR N-PSDB; AAR98106.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
PS  
PS Example 2; Figure 1; 41pp; English.

CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes.

XX Sequence 372 AA;

Query Match 100.0%; Score 666; DB 17; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNMQRRPFCRDNYTDLVAIQKAEIYELEKTLPSRSYWGIRKIGIW 60  
DB 39 WYHYSEKPMNMQRRPFCRDNYTDLVAIQKAEIYELEKTLPSRSYWGIRKIGIW 60  
QY 61 TWGNTKSLTEEAENWGDGEPNNKKNEDEVEIYIKRKKDAGKWNDDACHKLAALC 117  
DB 99 TWGNTKSLTEEAENWGDGEPNNKKNEDEVEIYIKRKKDAGKWNDDACHKLAALC 155

RESULT 24

ID AAR98133 standard; Protein; 372 AA.

XX AAR98133;

XX 01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

XX Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

XX Homo sapiens.

XX Key Location/Qualifiers

2

FT Region 1..37  
FT /label= Signal region.  
FT 38..154  
FT /label= Lectin domain.  
FT 159..192  
FT /label= EGF domain.  
FT 196..258  
FT /label= Complement binding repeat 1.  
FT 259..317  
FT /label= Complement binding repeat 2.  
FT 333..355  
FT /label= Transmembrane domain.  
FT 356..372  
FT /label= Cytoplasmic domain.

XX US514582-A.

XX 07-MAY-1996.

PD 23-FEB-1989; 89US-0315015.

PE 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX (GETH ) GENENTECH INC.

XX Capon DJ, Lasky LA;

XX WPI: 1996-238773/24.

XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
PS  
PS Disclosure; Page 19; 41pp; English.

CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a 209-Asn insertion.

XX Sequence 372 AA;

Query Match 100.0%; Score 666; DB 17; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2.3e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNMQRRPFCRDNYTDLVAIQKAEIYELEKTLPSRSYWGIRKIGIW 60  
DB 38 WYHYSEKPMNMQRRPFCRDNYTDLVAIQKAEIYELEKTLPSRSYWGIRKIGIW 97  
QY 61 TWGNTKSLTEEAENWGDGEPNNKKNEDEVEIYIKRKKDAGKWNDDACHKLAALC 117

[illegible]

CC	be used for preventing organ or graft rejection, for treating
CC	inflammatory disorders, e.g. rheumatoid arthritis or other
CC	autoimmune diseases, for controlling lymphoma metastasis and
CC	for treating conditions in which there is an accumulation of
CC	lymphocytes.
XX	
XX	Sequence    372 AA;
SQ	
Query Match	100.0%; Score 666; DB 19; Length 372;
Best Local Similarity	100.0%; Pred. NO. 2.3e-63;
Matches 117:	Conservative    0; Mismatches    0; Indels    0; Gaps    0;
QY	1 WTYHYSEKPMWQARRCRDNYTDLVAIQKAEELEYLEKTLPFSRSYYWGIRKGSIW 60
Dd	39 wtyhysekpmwqarrricrinydlvalqpkhaeeleyektlpfarsyywgirkigslw 98
QY	61 TWVGTSKSLTEBAENSGDEPNKKKNKEDCVETIYIKRNKAGKWNDACHKLKALC 117
Dd	99 twvgtskslteeaenwsgdepnkkknkedcveiyikrnkdagkwndachklkaalc 155
RESULT 26	
AAMW3264	
ID AAMW3264	standard; Protein; 372 AA.
XX	
AC AAW73264:	
DT	
XX 02-FEB-1999	(first entry)
XX	
DE Human lymphocyte homing receptor.	
XX	
KM Lymphocyte binding receptor; LHR: lymphocyte cell-surface glycoprotein;	
KW lymphocyte binding; endothelium; graft rejection; inflammation; therapy;	
KW arthritis; autoimmune disease; lymphoma metastasis;	
KW lymphocyte accumulation; human.	
XX	
OS Homo sapiens.	
XX	
PN US5840844-A.	
PD 24-NOV-1998.	
XX	
PF 10-AUG-1995;	95US-0513278.
XX	
PR 23-FEB-1989;	89US-0315015.
PR 31-OCT-1991;	91US-0786149.
PR 06-MAY-1993;	93US-0059029.
PR 10-AUG-1995;	95US-0513278.
XX	
PA (GETH ) GENENTECH INC.	
PA (REGC ) UNIV CALIFORNIA.	
XX	
P1 Lasky LA, Rosen SD, Singer MS, Stachel SE;	
XX	
DR WPI: 1999-034122/03.	
DR N-PsDB: AAV08321.	
XX	
PT Lymphocyte homing receptor polypeptides - useful for inhibiting	
PT lymphocyte binding to lymphoid endothelium	
XX	
PS Claim 1; Fig 1; 33pp; English.	
XX	
CC This sequence is the human lymphocyte homing receptor (LHR) of the	
CC invention. LHR is a lymphocyte cell-surface glycoprotein that mediates	
CC the binding of lymphocytes to the endothelium of lymphoid tissue. Soluble	
CC LHR polypeptides, lacking signal peptide (amino acids 1-38),	
CC transmembrane domain (amino acids 333-355) and cytoplasmic domain (amino	
CC acids 355-372), can be used therapeutically to compete with the normal	
CC binding of lymphocytes to lymphoid tissue and are especially useful for	
CC organ or graft rejection treatment protocols, for treating inflammations	
CC such as arthritis and other autoimmune diseases, for control of lymphoma	
CC metastasis and for treating conditions involving lymphocyte accumulation.	
CC	



PD 07-MAY-1996.

OS Homo sapiens.

PD 07-MAY-1996.

```

FH Key Location/Qualifiers
FT Region 347..368
FT /label= transmembrane
FT Modified-site 113..115
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 157..159
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 230..232
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 269..271
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 285..287
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 299..301
FT /label= N-linked_glycosylation
FT /note= "putative"
FT Modified-site 324..326
FT /label= N-linked_glycosylation
FT /note= "putative"
XX
XX W09201049-A.
XX
XX 23-JAN-1992.
XX
XX 15-JUL-1990; 90WO-US04986.
XX
XX 13-JUL-1990; 90US-0553759.
XX
XX (GEHO-) GEN HOSPITAL CORP.
XX
XX Seed B, Aruffo A, Amlot M;
XX
XX WPI; 1992-056864/07.
XX
XX N-PSDB; AAQ21184.
XX
XX New CD53 cell surface antigen and DNA encoding it - for
XX immuno-therapy and diagnosis of haematopoietic neoplasms, etc.
XX
XX Example 14; Page 106; 160pp; English.
XX
XX Two cDNA clones encoding Leu8 determinants were isolated from a
XX human T cell library using the rapid immunoselection cloning method
XX (see e.g. AAQ21164 for description of method). This protein sequence
XX was deduced from the larger insert. The hydrophobic putative
XX membrane-spanning domain is followed by several positively charged
XX residues resembling a cytoplasmic anchor sequence. The protein
XX is closely related to the murine Mel-14 homing receptor.
XX See AAQ25500 for the shorter insert and AAR22551 for the phospholipid
XX anchored form of Leu8 antigen.
XX
XX Sequence 385 AA:

```

```

Query Match 100.0%; Score 666; DB 13; Length 385;
Best Local Similarity 100.0%; Pred. No. 2,4e-63;
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 WTHYSEKPNWGRARFCNDNTDVAIONKAIEYLEKTLPSFSYWGIRKIGIWM 60
DB 52 WYHYSEKPNWGRARFCNDNTDVAIONKAIEYLEKTLPSFSYWGIRKIGIWM 111
OY 61 TWVGTNKSITLFEANWGDGEPNNKKNKEOCVEYIYIRKNDAGKWNDDACHKKAALC 117
DB 112 TWVGTNKSITLFEANWGDGEPNNKKNKEOCVEYIYIRKNDAGKWNDDACHKKAALC 168

```

```

RESULT 31
AAR32707

```

```

ID AAR32707 standard; Protein: 385 AA.
XX
XX AAR32707;
XX
XX 14-JUN-1993 (first entry)
XX
XX LAM-1 from pLAM-1.
XX
XX Leukocyte adhesion molecule; epitope; monoclonal antibody;
XX inflammation; autoimmune disease; shock; tissue damage; organ;
XX tissue; transplant; metastasis.
XX
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
XX Region 13..41
XX /note= "hydrophobic, possible signal sequence"
XX Region 346..362
XX /note= "hydrophobic, possible transmembrane sequence"
XX Modified-site 73..75
XX /note= "N-glycosylation site"
XX Modified-site 117..119
XX /note= "N-glycosylation site"
XX Modified-site 117..119
XX /note= "N-glycosylation site"
XX Modified-site 190..192
XX /note= "N-glycosylation site"
XX Modified-site 245..247
XX /note= "N-glycosylation site"
XX Modified-site 259..261
XX /note= "N-glycosylation site"
XX Modified-site 284..286
XX /note= "N-glycosylation site"
XX Modified-site 324..326
XX /note= "N-glycosylation site"
XX
XX W09302698-A.
XX
XX 18-FEB-1993.
XX
XX 23-JUL-1992; 92WO-US06127.
XX
XX 29-JUL-1991; 91US-0737092.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Spertini OG, Tedder TF;
XX
XX WPI; 1993-076177/09.
XX
XX N-PSDB; AAQ37304.
XX
XX Monoclonal antibodies to leukocyte adhesion molecule-1 (LAM-1) -
XX useful for treating inflammation, cancer, auto-immune diseases,
XX tissue damage and organ or tissue transplant rejection
XX
XX Disclosure; Fig 2; 121pp; English.
XX
XX B cell-specific cDNA was isolated from a human tonsil cDNA library
XX (ATCC #37546) using differential hybridisation with labelled cDNAs
XX from either B cell(RAJ1) RNA or T cell (HSB-2) RNA. Positive
XX plaques were isolated and cloned and the cDNA inserts subcloned into
XX pSP65. One of the 261 RAJ1+ HSB2- cDNA clones isolated, B125,
XX contained a 1.9 kb cDNA insert that hybridised with a 2.4 kb RNA
XX species found in several B cell lines. However B125 did not
XX hybridise with any of the other RAJ1+ HSB- clones or with mRNA from
XX several T cell lines. The B125 cDNA clone hybridised with a 2.3 kb
XX cDNA, termed pLAM-1 (prod. shown). The amino acid sequence of
XX LAM-1 indicates a structure typical of a membrane glycoprotein.
XX Monoclonal antibodies specific for LAM-1 may be prep. and are useful
XX in the treatment of inflammation, autoimmune disease, shock, tissue
XX damage, organ or tissue transplant rejection and to inhibit the
XX metastasis and homing of malignant cells which express the LAM-1
XX receptor protein. They can be used to monitor lamy-1 gene regulation,
XX

```



CC aggregation so prevent thrombosis. These bifunctional peptides can  
CC bind to 2 different selectin ligands so provide more effective  
CC targeting than monospecific compounds.  
XX  
SO Sequence 385 AA:

Query Match 100.0%; Score 666; DB 15; Length 385;  
Best Local Similarity 100.0%; Pred. No. 2.4e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 WYHYSEKPMNQRARRCRDNYTDVAIONKAIELEYLEKTLPEFSRSTYWGIRKIGTM 60  
Db 52 WYHYSEKPMNQRARRCRDNYTDVAIONKAIELEYLEKTLPEFSRSTYWGIRKIGTM 111  
QY 61 TWVGTKNSLTTEAENWGDGEPNNKKNKEDCVELYIRKNDAGKWNDDACHIKLKAALC 117  
Db 112 twvgtnkslteaenwgdgpnknkknkcdvcllylkrndagkwnddachiklkaalc 168

## RESULT 34

AAR91442  
ID AAR91442 standard; Protein: 385 AA.

XX AAR91442;

DT 31-OCT-1996 (first entry)

DE Human Leu8 antigen.

XX Cell surface antigen; cloning; immunoselection; immunotherapy;  
KW therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;  
XX antiInflammatory.

OS Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT Modified-site 113..115 /label= glycosylation\_site  
FT Modified-site 156..158 /label= glycosylation\_site  
FT Modified-site 229..231 /label= glycosylation\_site  
FT Modified-site 269..271 /label= glycosylation\_site  
FT Modified-site 285..287 /label= glycosylation\_site  
FT Modified-site 299..301 /label= glycosylation\_site  
FT Modified-site 324..326 /label= glycosylation\_site  
FT Domain /label= glycosylation\_site  
FT /label= Transmembrane\_domain  
FT 347..368

PN US5506126-A.

PD 09-APR-1996.

PF 25-FEB-1988; 88US-0160416.

XX 01-DEC-1992; 92US-0983647.

PR 25-FEB-1988; 88US-0160416.

PR 13-JUL-1989; 89US-0379076.

PR 18-OCT-1993; 93US-0553759.

XX (GEO) GEN HOSPITAL CORP.

PI Aruffo A, Seed B;

XX WPI; 1996-200279/20.

DR N-PSDB; AAT14723.

XX

PT Cloning of cDNA encoding cell surface antigen - useful for isolation  
PT of diagnostic and therapeutic proteins

PS Example 14; Column 71-74; 79pp; English.

XX 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were  
CC isolated from a human T-cell library using a novel immunoselection  
CC cloning method. The longer insert (AAT14723) contained 2,350  
CC residues, while the shorter lacked 436 internal residues. A major  
CC transcript of 2.4 kb was present in peripheral blood mononuclear  
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,  
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukaemic  
CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and  
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells  
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.  
CC Soluble forms of Leu8 can act as antiinflammatory agents by reducing  
CC lymphocyte migration.

XX Sequence 385 AA;

Query Match 100.0%; Score 666; DB 17; Length 385;  
Best Local Similarity 100.0%; Pred. No. 2.4e-63;  
Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 WYHYSEKPMNQRARRCRDNYTDVAIONKAIELEYLEKTLPEFSRSTYWGIRKIGTM 60  
Db 52 WYHYSEKPMNQRARRCRDNYTDVAIONKAIELEYLEKTLPEFSRSTYWGIRKIGTM 111  
QY 61 TWVGTKNSLTTEAENWGDGEPNNKKNKEDCVELYIRKNDAGKWNDDACHIKLKAALC 117  
Db 112 twvgtnkslteaenwgdgpnknkknkcdvcllylkrndagkwnddachiklkaalc 168

## RESULT 35

AAW21657  
ID AAW21657 standard; Protein: 385 AA.

XX AAW21657;

DT 30-SEP-1997 (first entry)

DE Human lymphocyte-associated cell surface protein.

XX Lymphocyte-associated cell surface protein; LAM-1; B lymphocyte;  
KW antibody; autoimmune disease; cancer.

XX Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT Peptide 1..51 /label= Sig-peptide  
FT Peptide 14..51 /label= Sig-peptide  
FT /label= Sig-peptide  
FT /note= "Signal peptide starting from alternative  
FT N-terminal Met"

FT Protein 52..385 /label= Mat\_protein

FT Domain 52..345 /label= Extracellular\_domain

FT Domain 346..362 /label= Transmembrane\_domain

FT Domain 363..385 /label= Intracellular\_domain

FT Modified-site 73..75 /label= glycosylation

FT /note= "putative N-glycosylation site"

FT Modified-site 117..119 /label= glycosylation

FT /note= "putative N-glycosylation site"

FT Modified-site 190..192 /label= glycosylation

FT /note= "putative N-glycosylation site"



FT	Modified-site	245..247	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	259..261	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	288..290	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
FT	Modified-site	324..326	/label= Glycosylation
FT		/note= "putative N-glycosylation site"	
XX			
PN	EP770680-A2.		
-XX			
PD	02-MAY-1997.		
XX			
PF	21-FEB-1990;	90EP-0301884.	
XX			
PR	21-FEB-1989;	89US-0313109.	
XX			
PA	(DAND ) DANA FARBER CANCER INST INC.		
XX			
PI	Tedder TF:		
XX			
DR	WPI: 1997-238140/22.		
DR	N-PSDB: AAT72270.		
XX			
PT	New antibodies to leukocyte adhesion molecule-1 - used for binding		
PT	LAM-1, e.g. for treating tissue damage, autoimmune disorders,		
PT	cancers or organ or tissue transplants		
XX			
PS	Claim 1; Page 6-8; 10pp: English.		
XX			
CC	Human lymphocyte-associated cell surface protein LAM-1 (AA021657)		
CC	contains domains homologous with binding domains of animal lectins,		
CC	growth factors and C3/C4 binding proteins. Its amino acid sequence		
CC	was deduced from a human tonsil cDNA clone (AAT72270). Together with		
CC	ELAM-1 and GMP-40, LAM-1 defines a new family of homologous		
CC	structures that are expressed by different cell lineages and which		
CC	can function as receptors in cellular interactions. LAM-1 can be		
CC	used as an antigen to raise antibodies capable of preventing		
CC	adhesion, migration or infiltration into tissues of cells		
CC	expressing the protein, or to identify antagonists useful for the		
CC	treatment of tissue damage, autoimmune disorders, cancer, or organ		
CC	or tissue transplants.		
XX			
XX	Sequence 385 AA:		
SO			
	Query Match	100.0%; Score 666; DB 18; Length 385;	
	Best Local Similarity	100.0%; Pred. No. 2.4e-63;	
	Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 WTYHSEKPMWQARRCRDNYTDLVAIQNKAIELEYLEKTLPFSSRYWIGIRKIGTW 60		
Db	52 wtyhysekpmwqarrircrdnytdlvalaigkaeyleylektpfarsywigirkigtw 111		
QY	61 TWVGKNSKLTLEAEWVGCGEPNNKKNKEDCVETIYIKRNKDAGKWDADCHIKLKAALC 117		
Db	112 twvgknskltleaeewvgdgepnnkknkedcveiyikrnkdagkwndadchiklkaalc 168		
	RESULT 36		
ID	AA080452		
XX	AA080452 standard; Protein; 385 AA.		
XX	AA080452;		
XX			
XX	07-JUN-1999 (first entry)		
XX			
DE	Human Leu8 antigen.		

KM		Leub; cell surface antigen; human; T lymphocyte; antiinflammatory;
XW		therapy.
XX	Homo sapiens.	
OS		
XX	Key	Location/Qualifiers
FH	Domain	347..368
FT		/note="transmembrane domain"
PN		
XX	US5830731-A.	
PD	03-NOV-1998.	
XX		
PF	21-MAY-1997;	970S-0861205.
XX		
PR	01-DEC-1992;	92US-0983647.
PP	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	13-JUL-1990;	90US-0553759.
PR	21-MAY-1997;	97US-0861205.
PA	(GEHO ) GEN.HOSPITAL CORP. *	
XX		
PI	Arafio A, Seed B:	
XX		
DR	WI: 1998-609251/51.	
DR	N-PSDB; AAU63460.	
XX	New cloning vector and polylinker - based on existing sequences for efficient cloning and expression of mammalian CDNA(s), especially human Lymphocyte antigenic sequences	
PT		
XX	Example 14; Column 67-70; 75bp; English.	
CC	This polypeptide comprises human LeuB antigen. Its amino acid sequence was deduced from the nucleotide sequence (see AAV63450) of a cDNA clone isolated from a T lymphocyte cDNA library using a novel method for cloning DNAs from mammalian expression libraries. The method is based on transient expression of an antigen in eukaryotic cells and physical selection of cells expressing the antigen by adhesion to an antibody-coated substrate. It is useful for the isolation and cloning of any protein which can be expressed and transported to the cell surface membrane of a eukaryotic cell, and was used to clone genes (see AAV63442-63) encoding cell surface antigens from mammalian lymphocytes (see AAU80440-55). The purified genes and proteins are useful for immunodiagnostic and immunotherapeutic applications, including the diagnosis and treatment of immune-mediated infections, diseases, and disorders of CC animals, including humans. The extracellular domain of LeuB, which mediates adhesion to specialised endothelial cells of lymph nodes, CC is quite specific in its recognition of the lectin ligand sulphated galactosyl ceramide (sulfatide). Modification of the specificity of this binding could serve to regulate the homing potential of resting T cells. Soluble forms of LeuB can act as anti-inflammatory agents by reducing lymphocyte migration.	
SQ	Sequence    385 AA:	
	Query Match                 100.0%; Score 666; DB 19; Length 385; Best Local Similarity      100.0%; Pred. No. 2.de-63; Matches 117; Conservative    0; Mismatches 0; Indels 0; Gaps  OY          1 WTVHHSEKMMNQRRARFRCDNYTDLVAIQNKAELFELEKTIPFSRYTWIGIRKGIM 60 Db          52 WCVHYHSKKPMNRQRITRFDYTDLVALAQNAAELEYELEKTPISRSYYWGIVGIRGIWM 111 OY          61 TWVGNNKSSTEEAEWNMGDESPNNKNKEPCVCETIYIKRNKDAGKWMDACGHRTAKALC 117 Db          112 TVWGNTKSLTEEAENWGDGPNNKRKEDGVCELYLKNKDAQGKWDDACHKIAALC 168	
RESULT 37		

AAW86199  
 ID AAW86199 standard; Protein; 385 AA.  
 AC AAW86199;  
 DT 10-MAY-1999 (first entry)  
 DE Human Leu8 antigen.  
 KW Leu8; cell surface antigen; human; CDNA library; T lymphocyte; antiinflammatory; therapy.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 347..368  
 FT /note="transmembrane domain"  
 XX US5849898-A.  
 XX 15-DEC-1998.  
 XX 07-JUN-1995; 95US-0485447.  
 XX 01-DEC-1992; 92US-0983647.  
 XX 25-FEB-1988; 88US-0160416.  
 XX 13-JUL-1989; 89US-0379076.  
 XX 23-MAR-1990; 90US-0498809.  
 XX 13-JUL-1990; 90US-0553759.  
 XX 07-JUN-1995; 95US-0485447.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PI Allen J, Amiot M, Arufo A, Camerini D, Laufer L;  
 PI Oquendo C, Seed B, Simmons D, Stamenkovic I, Stengelin S;  
 XX WPI: 1999-069813/06.  
 DR N-PSDB; AAW81217.  
 XX  
 PT cDNA encoding human CD40 antigen - useful for cloning cDNA encoding  
 PT cell surface antigens, constructing cDNA libraries, expression  
 PT vectors for expression in eukaryotic cells or their fragments  
 XX  
 PS Example 14; Column 67-70; 79pp; English.  
 XX  
 CC This polypeptide comprises human Leu8 antigen. Its amino acid  
 CC sequence was deduced from the nucleotide sequence (see AAW81217) of a  
 CC cDNA clone isolated from a T lymphocyte cDNA library using a  
 CC novel method for cloning cDNAs from mammalian expression libraries.  
 CC The method is based on transient expression of an antigen in  
 CC eukaryotic cells and physical selection of cells expressing the  
 CC antigen by adhesion to an antibody-coated substrate. It is useful  
 CC for the isolation and cloning of any protein which can be expressed  
 CC and transported to the cell surface membrane of a eukaryotic cell,  
 CC and was used to clone genes (see AAW81198-220) encoding cell surface  
 CC antigens such as CD1a, CD1b, CD1c, CD2, CD6, CD7, CD13, CD14, CD16,  
 CC CD19, CD20, CD22, CD26, CD27, CD28, CD31, CD32a, CD33b, CD33,  
 CC CD34, CD35, CD37, CD38, CD39, CD40, CD43, CD44, CD53, ICAM, LFA-3,  
 CC FCRIa, FCRIb, T118a and Leu8 (see AAW86188-62, AAW89151-52 and  
 CC AAW88451). CD40 cDNA (see AAW81198) is specifically claimed. The  
 CC extracellular domain of Leu8, which mediates adhesion to specialised  
 CC endothelial cells of lymph nodes, is quite specific in its recognition of  
 CC the lectin ligand sulphated galactosyl ceramide (sulfatide).  
 CC Modification of the specificity of this binding could serve to  
 CC regulate the homing potential of resting T cells. Soluble forms  
 CC of Leu8 can act as antiinflammatory agents by reducing  
 CC lymphocyte migration.  
 CC  
 XX Sequence 385 AA:

Query Match 100.0%; Score 666; DB 20; Length 385;  
 Best Local Similarity 100.0%; Pred. No. 2.4e-63;

Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 WTYHSEKPMNORARRCRDNYDVAIQNKATIELEKLPSSRSYTWIGIKIGIV 60  
 DB 52 WCYHYSEKPMNWGFAIRICRDYCDLVAIQKAELEYEKLPISRYGYVIGIKIGIW 111  
 QY 61 TWGVTNKSLEAEANWGDGEPNNKRNKEDCEYIYIKRNKDGAKNWDACRKLAAALC 117  
 DB 112 twgvtlnksleaeenwgdgpepnknknedcveiyikrnkdagkwnadackhkaalc 168  
 RESULT 38  
 ID AAY96138  
 AC AAY96138 standard; Protein; 385 AA.  
 DT 19-DEC-2000 (first entry)  
 DE Human T-cell specific Leu8 antigen.  
 KW Leu8; cell surface antigen; human; immunoselection; panning;  
 KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;  
 KW immune disorder; infection; asthma; immune-complex disease;  
 KW amyloidosis; multiple sclerosis; inflammation; antiinflammatory.  
 OS Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Domain 347..368  
 FT /label="Transmembrane\_domain"  
 XX US611093-A.  
 XX 29-AUG-2000.  
 XX 28-OCT-1998; 98US-0181612.  
 XX 01-DEC-1992; 92US-0983647.  
 XX 25-FEB-1988; 88US-0160416.  
 XX 13-JUL-1989; 89US-0379076.  
 XX 23-MAR-1990; 90US-0498809.  
 XX 13-JUL-1990; 90US-0553759.  
 PA (GEHO ) GEN HOSPITAL CORP.  
 PI Stamenkovic I, Seed B;  
 XX WPI: 2000-586382/55.  
 DR N-PSDB; AAA50596.  
 XX  
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,  
 PT useful for immunodiagnosis and immunotherapy of immune-mediated  
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic  
 PT diseases -  
 XX  
 PS Example 14; Column 69-70; 75pp; English.  
 XX  
 CC The present sequence is that of a T-cell specific Leu8 antigen, as  
 CC predicted from 1 of 2 clones (see AAA50596) isolated from a human  
 CC T-cell library by the method of the invention. A shorter Leu8  
 CC antigen is given in AAY96184. The method, designed to isolate cell  
 CC surface antigen (CSA) cDNAs, is based upon transient expression of  
 CC a CSA in eukaryotic cells and physical selection of cells expressing  
 CC the antigen by adhesion to (panning on) an antibody-coated substrate  
 CC such as a culture dish. CSA nucleic acids isolated by the method of  
 CC the invention, and the proteins they encode, are useful for  
 CC immunodiagnostic and immunotherapeutic applications, including the  
 CC diagnosis and treatment of immune-mediated infections, diseases, and  
 CC disorders in animals, including humans. These disorders include  
 CC asthma, immune-complex disease, amyloidosis, parasitic diseases or  
 CC multiple sclerosis. The ability to interfere with the binding of  
 CC Leu8-T-cells to antigen presenting cells, or the ability to cause

CC such binding to occur on surfaces other than lymphocyte cells, can  
 CC be useful in diagnostics and therapy. The level of activated Leu8-  
 CC T-cells relative to resting Leu8+ cells could serve as a measure of  
 CC immune response to a particular antigen. Modification of the  
 CC specificity of the extracellular domain of Leu8, which mediates  
 CC adhesion to specific endothelial cells of lymph nodes, could serve  
 CC to regulate the homing potential of resting T cells. Soluble forms  
 CC of Leu8 could act as antiinflammatory agents by reducing lymphocyte  
 CC migration.

XX  
 XX Sequence 385 AA;  
 SQ

Query Match 100.0%; Score 666; DB 21; Length 385;  
 Best Local Similarity 100.0%; Pred. No. 2,4e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNMORARFCRDNYTDLVAIONKAELEYLEKTLFESRSYWGIRKIGIM 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 52 wtyhsekpmnmqrrarfcridnytdlvaionkaeleylektlfrsrywgirkgilw 111  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 61 TWVGTNKSLTBEAKWNGDEPNKKKNEDEYIYIKRNKDAGKWDACCHKIKALC 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 112 twvgtnkslteeaenwgdepnknknedcveilykrnkdgakwndachikikaalc 168  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 39  
 AAU02447  
 ID AAU02447 standard; Protein: 405 AA.  
 AC AAU02447;  
 XX  
 XX 29-AUG-2001 (first entry)  
 DT  
 DE Human T-lymphocyte specific antigen Leu8 polypeptide.  
 XX  
 XX Human: T-lymphocyte specific antigen; immune-mediated disease; Leu8;  
 KW infection; immune deficiency disorder; hypersensitivity; inflammation;  
 KW systemic lupus erythematosus; platelet disorder; rheumatoid arthritis;  
 KW transplant rejection; asthma.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Modified-site 113..115  
 FT Modified-site /note= "N-linked glycosylation site"  
 FT Modified-site 157..159  
 FT Modified-site /note= "N-linked glycosylation site"  
 FT Modified-site 230..232  
 FT Modified-site /note= "N-linked glycosylation site"  
 FT Modified-site 269..271  
 FT Modified-site /note= "N-linked glycosylation site"  
 FT Modified-site 285..287  
 FT Modified-site /note= "N-linked glycosylation site"  
 FT Modified-site 324..326  
 FT Modified-site /note= "Asn is N-glycosylated"  
 FT MISC-difference 344..385  
 FT /note= "These amino acids are present in the longer  
 FT form but are absent in the short form of Leu8"  
 FT  
 FT Domain 347..368  
 FT /label= Transmembrane\_domain  
 XX  
 XX US6218525-B1.  
 FN  
 XX 17-APR-2001.  
 PD  
 XX 01-DEC-1992; 92US-0983647.  
 PE  
 XX 25-FEB-1988; 88US-0160416.  
 PR 13-JUL-1989; 89US-0379076.  
 PR 13-JUL-1990; 90US-0553759.  
 XX  
 PA (GENO ) GEN HOSPITAL CORP.

XX  
 PI Seed B, Aruffo A, Simmons D;  
 XX  
 XX WPI: 2001-289848/30.  
 DR N-PSDB: AAS03190.  
 XX  
 XX New recombinant DNA encoding CD28 useful for diagnosing and treating  
 PT immune-mediated diseases, infections or disorders, e.g. systemic lupus  
 PT erythematosus, asthma, transplant rejection, rheumatoid arthritis  
 PT  
 XX  
 XX Example 14: Column 63-66; 72pp; English.  
 PS  
 XX  
 XX The present sequence represents human T-lymphocyte specific antigen  
 CC Leu8 encoded by a long leu8 cDNA clone. The shorter cDNA clone lacks 436  
 CC bases and consequently lacks residues 344-385 given here. Various human  
 CC lymphocyte cell surface antigen polypeptide sequences  
 CC (AAU02435-AAU02452) are described in the present invention. The invention  
 CC relates to a novel method of cloning cDNA encoding cell surface antigens  
 CC and efficient construction of cDNA libraries. Also described are 2  
 CC expression vectors (AAS03171, AAS03174) which provide high level  
 CC expression in eukaryotic host cells. A genetically engineered cDNA  
 CC sequence encoding the CD28 amino acid extracellular domain sequence  
 CC (amino acids 1-134 given in AAU02437) and/or comprising nucleotides  
 CC 100-759, 154-555 or 154-759 of the CD28 cDNA sequence (AAS03175) is also  
 CC new. The purified genes and proteins are useful for immunodiagnostic and  
 CC immunotherapeutic applications, such as in the diagnosis and treatment of  
 CC immune-mediated diseases, infections or disorders in animals and humans.  
 CC Such diseases include immune deficiency diseases, diseases of immediate  
 CC type of hypersensitivity, asthma, hypersensitivity pneumonitis, systemic  
 CC lupus erythematosus, rheumatoid arthritis, acute and chronic  
 CC inflammation, platelet disorders, plasma and other cell neoplasms,  
 CC parasitic diseases, multiple sclerosis, Guillain-Barre syndrome and  
 CC tissue and organ transplant rejection. The sequences can also be used to  
 CC identify, isolate and purify other antibodies and antigens.  
 CC  
 XX  
 XX Sequence 405 AA;  
 SQ

Query Match 100.0%; Score 666; DB 22; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 2,3e-63;  
 Matches 117; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 WTYHSEKPMNMORARFCRDNYTDLVAIONKAELEYLEKTLFESRSYWGIRKIGIM 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 52 wtyhsekpmnmqrrarfcridnytdlvaionkaeleylektlfrsrywgirkgilw 111  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

OY 61 TWVGTNKSLTBEAKWNGDEPNKKKNEDEYIYIKRNKDAGKWDACCHKIKALC 117  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 112 twvgtnkslteeaenwgdepnknknedcveilykrnkdgakwndachikikaalc 168  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

RESULT 40  
 AAR98110  
 ID AAR98110 standard; Protein: 371 AA.  
 AC AAR98110;  
 XX  
 XX 01-NOV-1996 (first entry)  
 DT  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 XX  
 XX  
 XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 KW  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Region 1..37  
 FT Domain /label= Signal region.  
 FT 38..154  
 FT /label= Lectin domain.

FT Domain 159..192  
 FT /label= BGF domain.  
 FT Binding-site 196..257  
 FT /label= Complement binding repeat 1.  
 FT Binding-site 258..316  
 FT /label= Complement binding repeat 2.  
 FT Domain 332..354  
 FT /label= Transmembrane domain.  
 FT Domain 355..371  
 FT /label= Cytoplasmic domain.

PN US5514582-A.

XX 07-MAY-1996.

PD 23-FEB-1989;

XX 89US-0315015.

PR 22-NOV-1989;

PR 23-FEB-1989;

PR 16-DEC-1991;

PR 08-DEC-1992;

PR 21-JAN-1994;

XX 94US-0185670.

XX (GETH ) GENENTECH INC.

XX Capon DJ, Lasky LA;

PI WPI; 1996-238773/24.

XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand

PT binding site of a receptor fused to Ig constant region - useful for

PT diagnosis and treatment e.g. of inflammation

XX Disclosure; Page 19; 41pp; English.

PS A hybrid immunoglobulin chain comprising the ligand binding site of

XX a single transmembrane receptor without an active transmembrane

CC region; fused at its C-terminus with the N-terminus of an

CC immunoglobulin constant region. The receptor is not a member of the

CC immunoglobulin super family, nor a multiple subunit polypeptide

CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the

CC effector functions of immunoglobulin and can bind to and/or activate

CC more than one ligand. It can be used diagnostically for the in

CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.

CC Typical applications are as antiviral, neuromodulating and

CC immunomodulating agents, or as modulators of cell adhesion (e.g. in

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)

CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region

CC facilitates recovery, improves aqueous solubility and removes

CC potentially immunogenic epitopes. Variants of the human lymphocyte

CC cell surface glycoprotein described in AAR98106 are given in

CC AAR98109-R98135. This variant contains an Ala71Ser substitution.

XX SQ Sequence 371 AA;

XX Query Match 99.5%; Score 663; DB 17; Length 371;

XX Best Local Similarity 99.1%; Pred. No. 4.8e-63;

XX Matches 116; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 WTYHYSKPKPMQRRARFCNDYTDVAIONKAEIEYLEKTLPFSSRSYWIIGIRKIGIW 60

Db 38 WYHYSEKPMNQRRARFCNDYTDVAIONKAEIEYLEKTLPFSSRSYWIIGIRKIGIW 97

QY 61 TWVGTKNSLITEAENMGDGEPPNKKKKEDCEVEIYIKRNKDAGKWNDDACHLKAALC 117

Db 98 TWVGTKNSLITEAENMGDGEPPNKKKKEDCEVEIYIKRNKDAGKWNDDACHLKAALC 154

Search completed: September 7, 2002, 10:14:07  
 Job time: 265 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:17 : Search time 61.04 Seconds  
(without alignments)  
96.360 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193  
Perfect score: 215  
Sequence: 1 CQPMSCSGHGECEIINNHTNCNDVGYGPGQDL 34

Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP\_invertebrate:\*  
6: SP\_mammal:\*  
7: SP\_mhc:\*  
8: SP\_organelle:\*  
9: SP\_phage:\*  
10: SP\_plant:\*  
11: SP\_protent:\*  
12: SP\_virus:\*  
13: SP\_unclassified:\*  
14: SP\_virus:\*  
15: SP\_virus:\*  
16: SP\_bacteriap:\*  
17: SP\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	205	95.3	385	4 Q9UJ43	Q9UJ43 homo sapien
2	181	84.2	376	6 Q28629	Q28629 oryctolagus
3	159	74.0	372	11 Q63762	Q63762 rattus norv
4	131	60.9	609	6 Q9GLFO	Q9GLFO canis faml
5	127	59.1	484	6 Q95LG2	Q95LG2 ovis aries
6	125	58.1	649	6 Q28657	Q28657 oryctolagus
7	125	58.1	482	6 Q28982	Q28982 sus scrofa
8	125	58.1	616	4 Q95509	Q95509 homo sapien
9	125	58.1	740	4 Q95507	Q95507 homo sapien
10	125	58.1	740	4 Q95508	Q95508 homo sapien
11	123	57.2	610	6 Q95IG1	Q95IG1 equus caball
12	122	56.7	646	6 Q28097	Q28097 sus scrofa
13	121	56.3	485	6 Q95LG3	Q95LG3 odocoileus
14	119	55.3	754	6 Q28290	Q28290 canis faml
15	98	45.6	1551	5 Q9NGV4	Q9NGV4 drosophila
16	98	45.6	3396	5 Q9YV55	Q9YV55 drosophila

17	95	44.2	710	5 Q93563	Q93563 caenorhabd
18	94	43.7	112	4 Q9H557	Q9H557 homo sapien
19	93	43.3	642	13 P79941	P79941 xenopus lae
20	93	43.3	723	4 Q9UJ43	Q9UJ43 homo sapien
21	93	43.3	723	4 Q9NU41	Q9NU41 homo sapien
22	91	42.3	2656	5 Q9GNUM3	Q9GNUM3 paracentrot
23	90	41.9	721	13 Q91902	Q91902 xenopus lae
24	90	41.9	2531	5 Q16004	Q16004 lytechinus
25	90	41.9	3767	5 Q9UA13	Q9UA13 caenorhabd
26	89.5	41.6	2524	5 Q9GPA5	Q9GPA5 branchiosto
27	89	41.4	1203	11 Q06008	Q06008 mus musculu
28	89	41.4	2281	4 Q9UPL3	Q9UPL3 homo sapien
29	89	41.4	2319	11 Q9UPL7	Q9UPL7 rattus norv
30	89	41.4	2321	4 Q9Y618	Q9Y618 homo sapien
31	89	41.4	2321	4 Q9UM47	Q9UM47 homo sapien
32	89	41.4	2470	11 Q95516	Q95516 mus musculu
33	89	41.4	2471	4 Q04721	Q04721 homo sapien
34	89	41.4	2471	4 Q9H240	Q9H240 homo sapien
35	89	41.4	2471	11 Q9OW30	Q9OW30 rattus sp..
36	89	41.4	2653	5 Q25253	Q25253 lucilia cup
37	88.5	41.2	3209	13 Q93574	Q93574 gallus gall
38	88	40.9	529	5 Q25058	Q25058 heliocidari
39	88	40.9	562	5 Q9VQ12	Q9VQ12 drosophila
40	88	40.9	585	5 Q9U0E2	Q9U0E2 tribolium c
41	88	40.9	721	5 Q95YGO	Q95YGO clona savig
42	87	40.5	156	5 Q26661	Q26661 strongyloce
43	87	40.5	174	5 Q9NA32	Q9NA32 caenorhabd
44	87	40.5	460	11 Q99PC6	Q99PC6 mus musculu
45	87	40.5	460	11 Q91WN8	Q91WN8 mus musculu

## ALIGNMENTS

RESULT 1  
ID Q9UJ43 PRELIMINARY: PRT: 385 AA.  
AC Q9UJ43:  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE L-SELECTIN PRECURSOR.  
GN L-SELECTIN.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);  
RA Fieger C.B.;  
RC Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);  
RA Fieger C.B.;  
RC Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.  
RL EMBL; AJ246000; CAB5488.1; -.  
DR HSSP; P1451; 1KTB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; Lectin.C.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR Pfam; PF00008; EGF\_1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR PRINTS; PR00132; SELECTIN.  
DR SMART; SM00032; CCP\_2.  
DR SMART; SM00034; CUECT; 1.  
DR SMART; SM00181; EGF\_1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.

DR PRINTS; PR00343; SELECTIN  
DR SMART; SM00032; CCP; 6.  
DR SMART; SM00034; SELECTIN

DR SMART; SM00181; EGF; 4.  
 DR SMART; PS00001; EGF\_1like; 1.  
 DR PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 KW Lectin; Selectin.  
 SQ SEQUENCE 609 AA; 66073 MW; 41862D1F4D23881F CRC64;

Query Match 60.9%; Score 131; DB 6; Length 609;  
 Best Local Similarity 60.6%; Pred. No. 5.4e-10;  
 Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33  
 |||||  
 DB 144 CPTSCSGHCEVEITNNHTCNCVPGFRLRCE 176

RESULT 5  
 O95LG2 PRELIMINARY; PRT; 484 AA.

AC 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-21421234; PubMed-11529941;  
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
 RA MacLachlan N.J.;  
 RT "Characterization of equine E-selectin."  
 RL Immunology 103:498-504(2001).  
 DR EMBL: AF307971; AAK4871.1; -.  
 KW Lectin; Selectin.  
 SQ SEQUENCE 484 AA; 52951 MW; 617930C1C2F47B44 CRC64;

Query Match 59.1%; Score 127; DB 6; Length 484;  
 Best Local Similarity 60.6%; Pred. No. 1.5e-09;  
 Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33  
 |||||  
 DB 143 CPTSCSGHCEVEITNNHTCNCVPGFRLRCE 175

RESULT 6  
 O28657 PRELIMINARY; PRT; 649 AA.

AC 028657;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE P-SELECTIN.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Vora D.K., Fang Z., Liva S.M., Parham F., Watson A.D., Drake T.A.,  
 RA Terfio M.C., Berlinger J.A.;  
 RT "Induction of P-selectin by MM-IDL and its role in human  
 RT atherosclerosis."  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Warden C.H.;

RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; U39446; AA81385.1; -.  
 DR HSSP; P16109; IFSB.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR InterPro; IPR002396; Selectin.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00059; Lectin\_C; 1.  
 DR Pfam; PF00084; Sush1; 6.  
 DR PRINTS; PR00343; SELECTIN.  
 DR SMART; SM00032; CCP; 6.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
 DR PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 KW EGF-like domain; glycoprotein; Lectin; Selectin.  
 SQ SEQUENCE 649 AA; 71755 MW; ECCD8C847B84BC31 CRC64;

Query Match 59.1%; Score 127; DB 6; Length 649;  
 Best Local Similarity 57.6%; Pred. No. 2.1e-09;  
 Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

OY 1 COPMSCSGHCEVEITNNHTCNCVDGYGPOCQ 33  
 |||||  
 DB 163 CODMSCSKQECIETIGNYTCSCYPGYGPECE 195

RESULT 7  
 O28982 PRELIMINARY; PRT; 482 AA.

AC 028982;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97075911; PubMed-8918234;  
 RA Winkler H., Brostjan C., Szizmadia V., Natarajan G., Anrather J.,  
 RA Bach F.H.;  
 RT "The Intron-exon structure of the porcine E-selectin-encoding gene."  
 RL Gene 176:67-72(1996).  
 DR EMBL; U37521; AAC48680.1; -.  
 DR HSSP; P16581; IESL.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR InterPro; IPR002396; Selectin.  
 DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam; PF00008; EGF; 1.  
 DR Pfam; PF00059; Lectin\_C; 1.  
 DR Pfam; PF00084; Sush1; 4.  
 DR PRINTS; PR00343; SELECTIN.  
 DR SMART; SM00032; CCP; 4.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; 1.  
 DR PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 KW EGF-like domain; glycoprotein; Lectin; Selectin.  
 SQ SEQUENCE 482 AA; 52341 MW; 97DC5D70BF115944 CRC64;

Query Match 58.1%; Score 125; DB 6; Length 482;  
 Best Local Similarity 60.6%; Pred. No. 2.9e-09;

	Matches	20; Conservative	4; Mismatches	9; Indels	0; Gaps
QY	1	CQPMSCGHGECVELIINHTCMCDVGYGPOCQ	33		
Db	144	CTPTSCSGHGECIETINSTCCQYGFEGLOCE	176		

Query Match	Best Local Similarity	58.1%	Score 125	DB 4	Length 616
Matches 19	Conservative 6	Mismatches 8	Indels 0	Gaps 0	
0y 1 CPMSCSGHGCEVIEINNHTCNDVGYGBOQC 33	11 111 111 1 1 111 1 1 111 1 1				
Db 163 CDMSCSGKQEGCELTETIGNYTCSCYPGYGPECE 195					
RESULT 9	PRELIMINARY	PRT	740 AA		
ID 095507	PRELIMINARY	PRT	740 AA		
AC 095507					
DT 01-MAY-1999 (TrEMBLrel. 10, Created)					
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)					
DT 01-DEC-2001 (TrEMBLrel. 19, last annotation update)					
DE D1780M13.1.1 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN DB2, GMP140) (ISOFORM 1) (FRAGMENT)).					
DE CB62, GMP140) (ISOFORM 1) (FRAGMENT)).					
DE SELP.					
OS Homo sapiens (Human).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.					
OC NCBI_TaxID=9606;					
RN [1]					
RP SEQUENCE FROM N.A.					
RA Howden P.;					
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AL022146; CAA18144.1; -.					
DR HSSP; P16109; 1KID.					
DR InterPro; IPR000561; EGF-like.					
DR InterPro; IPR001304; lectin_c.					
DR InterPro; IPR002396; selectin.					
DR InterPro; IPR000436; Sush1_SCR_CCP.					
DR Pfam; PF00008; EGF; 1.					
DR Pfam; PF00059; lectin_c; 1.					
DR Pfam; PF00084; sush1; 6.					
DR PRINTS; PR00343; SELECTIN.					
DR SMART; SM00032; CCP; 6.					
DR SMART; SM00034; CLECT; 1.					
DR SMART; SM00181; EGF; 1.					
DR PROSITE; PS00615; C_TYPE_LECTIN_1; 1.					
DR PROSITE; PS50041; C_TYPE_LECTIN_2; 1.					
DR PROSITE; PS00022; EGF_1; UNKNOWN_1.					
DR PROSITE; PS01186; EGF_2; 1.					
KW EGF-like domain; Glycoprotein.					
FT NON_TER 616 616					
SO SEQUENCE 616 AA; 67736 MW; 35CD4BFADE61D274 CRC64;					

RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases  
 DR EMBL: AI022146; CAA18142.1; -.  
 DR HSSP: P16109; 1KJD.  
 DR Interpro: IPR000561; EGF-like.  
 DR Interpro: IPR001304; Lectin\_C.  
 DR Interpro: IPR002396; Selectin.  
 DR Interpro: IPR000436; Sushi\_NCCP.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; Lectin\_C\_1.  
 DR Pfam: PF00084; sushi\_8.  
 DR PRINTS: PRO0343; SELECTIN.  
 DR SMART: SM00032; CCP; 8.  
 DR SMART: SM00034; CLECT\_1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE: PS50041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW EGF-like domain; Glycoprotein.  
 FT 740 740  
 SQ NON\_TER  
 SEQUENCE 740 AA; 81095 MW; 3B5F70A45B1A3CD4 CRC64;

Query Match	58.1%;	Score 125;	DB 4;	Length 740;
Best Local Similarity	57.68;	Pred. No. 4.5e-09;		
Matches 19;	Conservative 6;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCDVGYGPOCQ 33  
||| |||:| |:|:| |:|:|:|:  
Db 163 CQPMSCSKQGECLLETIGNYTCSCYPGFYGPECE 195

RESULT	10			
095508				
ID	095508	PRELIMINARY:	PRT:	740 AA.
AC	095508;			
DT	01-MAY-1999 (TREMBLrel. 10, Created)			
DT	01-MAY-1999 (TREMBLrel. 10, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	DJ780M13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN			
DE	CD62, GMP140)) (ISOFORM 2) (FRAGMENT).			
GN	SELP.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Howden P.;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AL022146; CAA18143.1; -;			
DR	HSSP: P16109; IKD.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001304; lectin_C.			
DR	InterPro: IPR002396; Selectin.			
DR	InterPro: IPR00436; Sush1_SCR_CCP.			
DR	Pfam: PF00008; EGF_1.			
DR	Pfam: PF00059; lectin_C_1.			
DR	Pfam: PF00084; sush1_8.			
DR	PRINTS: PR00343; SELECTIN.			
DR	SMART: SM00032; CCP_8.			
DR	SMART: SM00034; CLECT_1.			
DR	SMART: SM00181; EGF_1.			
DR	PROSITE: PS00615; C_TYPE_LECTIN_1; 1.			
DR	PROSITE: PS00641; C_TYPE_LECTIN_2; 1.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE: PS01186; EGF_2; 1.			
KW	EGF-like domain; Glycoprotein.			
FT	NON_TER	740	740	
SEQUENCE	740 AA; 81390 MM; 1D2E35ED93745CE CRC64;			

Query Match 58.18; Score 125; DB 4; Length 740;



Best Local Similarity 57.6%; Pred. No. 4,5e-09;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33  
DB 163 CQSTSCSKGCECTETIGNYTCSCYSGFYGPCE 195

## RESULT 11

QY 095LG1 PRELIMINARY; PRT; 610 AA.  
AC 095LG1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE E-SELECTIN.  
OS Equus caballus (Horse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
OX NCBI\_TaxID=9796;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21421234; PubMed-11529941;  
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
RA MacLachlan N.J.;  
RT "Characterization of equine E-selectin."  
RL Immunology 103:498-504(2001).  
DR EMBL; AF307972; AAK48712.1; -.  
KW lectin; Selectin.  
SQ SEQUENCE 610 AA; 66191 MW; F9D3DED12C445382 CRC64;

Query Match 57.2%; Score 123; DB 6; Length 610;  
Best Local Similarity 60.6%; Pred. No. 6,9e-09;

Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33  
DB 142 CHTSCSGHGECEIINNHTCNCNDVGYGPOCQ 174

## RESULT 12

QY 029097 PRELIMINARY; PRT; 646 AA.  
AC 029097;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE P-SELECTIN PRECURSOR.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-AORTA;  
RA Rollins S.A., Johnson K.K., Birks C.W., Mattis L.A., Rother R.P.;  
RL Submitted (Oct-1995) to the EMBL/Genbank/DBO databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-AORTA;  
RX MEDLINE-20171534; PubMed-10706724;  
RA Stocker C.J., Sugars K.L., Harari O.A., Landis R.C., Morley B.J.,  
RA Haekard D.O.;  
RT "TNF-alpha, IL-4, and IFN-gamma regulate differential expression of P-  
and E-selectin expression by porcine aortic endothelial cells."  
RL J. Immunol. 164:3309-3315(2000).  
DR EMBL; L39075; AAA79007.1; -.  
DR EMBL; AF163766; AAF43272.1; -.  
DR HSPF; P16109; IFSB.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001304; lectin-C.  
DR InterPro; IPR002396; Selectin.

DR InterPro; IPR000436; Sush1\_SCR\_CCP.

DR Pfam; PF00008; EGF\_1.  
DR Pfam; PF00059; lectin\_C\_1.  
DR Pfam; PF00084; sush1\_6.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 6.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C\_TYPE\_LLECTIN\_1; UNKNOWN\_1.  
DR PROSITE; PS00041; C\_TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.  
FT SIGNAL 1 41 POTENTIAL.  
FT CHAIN 42 646 P-SELECTIN.  
SQ SEQUENCE 646 AA; 71127 MW; 3863F4FE09F0BB6 CRC64;

Query Match 56.7%; Score 122; DB 6; Length 646;  
Best Local Similarity 57.6%; Pred. No. 1e-08;

Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33  
DB 163 CQSTSCSKGCECTETIGNYTCSCYSGFYGPCE 195

## RESULT 13

QY 095LG3 PRELIMINARY; PRT; 485 AA.  
AC 095LG3;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE E-SELECTIN.  
OS Odocoileus hemionus (Mule deer) (Black-tailed deer).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
OC Cervidae; Odocoileinae; Odocoileus.  
OX NCBI\_TaxID=9872;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-21421234; PubMed-11529941;  
RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
RA MacLachlan N.J.;  
RT "Characterization of equine E-selectin."  
RL Immunology 103:498-504(2001).  
DR EMBL; AF307970; AAK48710.1; -.  
KW lectin; Selectin.  
SQ SEQUENCE 485 AA; 53247 MW; 69959199EAFE9980 CRC64;

Query Match 56.3%; Score 121; DB 6; Length 485;  
Best Local Similarity 57.6%; Pred. No. 1e-08;

Matches 19; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHTCNCNDVGYGPOCQ 33  
DB 144 CHTPCSSHGECEIINNHTCNCNDVGYGPOCQ 176

## RESULT 14

QY 028290 PRELIMINARY; PRT; 754 AA.  
AC 028290;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DE CELL ADHESION MOLECULE PRECURSOR (FRAGMENT).  
GN GMP140.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OK NCBI\_TaxID=9615;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC TISSUE=SPLEEN;  
RA Manning A.M., Sanders W.E.Jr., Kukielka G.L., Dore M.,  
RA Rosenbloom C.L., Hawkins H.L., Michael L.H., Ertman M.L., Smith C.W.,  
RA Beaudet A.L., Anderson D.C.,  
RT "Molecular cloning of canine GMP140 and studies of expression in a  
RT model of myocardial ischemia/reperfusion."  
RL Submitted (MAR-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL: M88170; AAA63789.1; -  
DR HSSP: P16109; 1FSB  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin\_c.  
DR InterPro: IPR002396; Selectin.  
DR Pfam: PF000436; Sushi\_SCR\_CCP.  
DR Pfam: PF00059; lectin\_c; 1.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP; 8.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS00615; C-TYPE LECTIN\_1; 1.  
DR PROSITE: PS50041; C-TYPE LECTIN\_2; 1.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR EGF-like domain; Glycoprotein; Signal.  
FT NON\_TER 1  
FT SIGNAL <1 13 POTENTIAL.  
FT CHAIN 14 754 CELL ADHESION MOLECULE.  
FT SEQUENCE 754 AA; 82303 MW; F0438BEA5521E773 CRC64;

Query Match 55.3%; Score 119; DB 6; Length 754;  
Best Local Similarity 54.5%; Pred. No. 3.1e-08;  
Matches 18; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
Qy 1 CQPMSCSGHCEVETIINHTNCNCDVGYGPQCQ 33  
Db 148 CQPMSCSKQEGCEITETIGMTCSPFGYSGCE 180

RESULT 15  
Q9NGV4 PRELIMINARY; PRT; 1551 AA.  
AC Q9NGV4; 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE SP1070.  
GN SP1070 OR CG9138.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ENDOPLASMIC RETICULUM;  
RA Serrano T.L., Pendleton J.D., Rubin G.M.;  
RT "A reverse genetic screen for genes involved in Drosophila  
RT development."  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF239608; AAF63500.1; -  
DR HSSP: P00740; 1EDM.  
DR FlyBase: FBgn0031879; SP1070.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF-2.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF-IT.  
DR InterPro: IPR003410; HYR.  
DR InterPro: IPR002049; Laminin\_EGF.

DR InterPro: IPR001791; Laminin\_G.  
DR InterPro: IPR002172; LDL\_recept\_A.  
DR Pfam: PF00008; EGF; 16.  
DR Pfam: PF02494; HYR; 1.  
DR PRINTS: PR00010; EGFBL00D.  
DR PRINTS: PR00011; EGFBLAMININ.  
DR SMART: SM00179; EGF\_CA; 6.  
DR SMART: SM00001; EGF-like; 9.  
DR SMART: SM00282; Lang; 1.  
DR SMART: SM00182; LDua; 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_9.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_15.  
DR PROSITE: PS01186; EGF\_2; 12.  
DR PROSITE: PS01187; EGF\_CA; 5.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.  
SQ SEQUENCE 1551 AA; 167816 MW; A97EA229E9384F31 CRC64;

Query Match 45.6%; Score 98; DB 5; Length 1551;  
Best Local Similarity 39.4%; Pred. No. 5.3e-05;  
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;  
Qy 1 CQPMSCSGHCEVETIINHTNCNCDVGYGPQCQ 33  
Db 251 CENPRCANEGTCVDLVGYSCNCEPGYTKNQC 283

RESULT 16  
Q9VM55 PRELIMINARY; PRT; 3396 AA.  
AC Q9VM55; 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CG9138 PROTEIN.  
GN SP1070 OR CG9138.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.B., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson G., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
RA Jalaal M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laspo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Splier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zhang X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Zhu H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
CC -1 SIMILARITY: CONTAINS 3 CUB DOMAINS.  
DR EMBL: AE003615; AAF52472.1; -.  
DR HSSP: P00740; 1EDM.  
DR FlyBase: FBgn0031879; SPL070.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000859; CUB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF-2.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_IT.  
DR InterPro: IPR000421; FAS58\_C.  
DR InterPro: IPR001092; HLH\_dim.  
DR InterPro: IPR003410; HYR.  
DR InterPro: IPR001791; Laminin\_G.  
DR InterPro: IPR002172; LDL\_recept\_A.  
DR InterPro: IPR000436; Sushl\_SCR\_CCP.  
DR InterPro: IPR001368; TNFR\_C6.  
DR Pfam: PF00431; CUB; 3.  
DR Pfam: PF00008; EGF; 16.  
DR Pfam: PF00754; F5\_F8\_type\_C; 2.  
DR Pfam: PF02494; HYR; 3.  
DR Pfam: PF00057; Idl\_recept\_a; 1.  
DR Pfam: PF00084; sushl; 7.  
DR PRINTS: PR00010; EGFBLDOD.  
DR SMART: SM00032; CCP; 8.  
DR SMART: SM00042; CUB; 3.  
DR SMART: SM00179; EGF\_CA; 8.  
DR SMART: SM00001; EGF-like; 9.  
DR SMART: SM00231; FAS58; 2.  
DR SMART: SM00282; LamG; 1.  
DR SMART: SM00192; LDLa; 1.  
DR SMART: SM00208; TNFR; 2.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 11.  
DR PROSITE: PS01180; CUB; 6.  
DR PROSITE: PS00022; EGF\_1; 15.  
DR PROSITE: PS01186; EGF\_2; 13.  
DR PROSITE: PS01187; EGF\_CA; 7.  
DR PROSITE: PS01285; FAS58\_1; 1.  
DR PROSITE: PS01209; LDLa\_1; 1.  
DR PROSITE: PS00068; LDLa\_2; 2.  
KW Calcium-binding; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.  
SQ SEQUENCE 3396 AA; 369389 MW; E618E9ACEA13E0E5 CRC64;

Query Match 45.6%; Score 98; DB 5; Length 3396;  
Best Local Similarity 39.4%; Pred. No. 0.00012;  
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINHTNCNDVGYGPOCQ 33  
Db 2092 CENRPCANEGTCVLDVDSYSCNCEPGYTGKNQ 2124

RESULT 17  
Q93563 PRELIMINARY; PRT; 710 AA.

AC 093563;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE F25D7.5 PROTEIN.  
GN F25D7.5.  
OS *Caenorhabditis elegans*.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodierinae; *Caenorhabditis*.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA McMurry A.A.;  
RL Submitted (Aug-1996) to the EMBL/Genbank/DBJ databases.  
RP [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode *C. elegans*: A platform for  
RT investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL: Z78418; CAB01698.1; -.  
DR InterPro: IPR000561; EGF-like.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE: PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein.  
SQ SEQUENCE 710 AA; 80783 MW; 4F5DF8112C5D558 CRC64;

Query Match 44.2%; Score 95; DB 5; Length 710;  
Best Local Similarity 38.2%; Pred. No. 6.2e-05;  
Matches 13; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINHTNCNDVGYGPOCQ 34  
Db 148 CPTTCNGHCKCYDVEDVKDCQWYGEHCEV 181

RESULT 18  
Q9H557 PRELIMINARY; PRT; 112 AA.

AC Q9H557;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE DJ303F19.1 (SIMILAR TO NOTCH (DROSOPHILA) HOMOLOG) (FRAGMENT).  
GN DJ303F19.1.  
OS *Homo sapiens* (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Williams S.;  
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AL137007; CAC10575.1; -.  
DR HSSP: P00749; 1URK.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF-2.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_IT.  
DR PRINTS: PR00010; EGFBLDOD.  
DR SMART: SM00181; EGF; 3.  
DR SMART: SM00179; EGF\_CA; 1.  
DR SMART: SM00001; EGF-like; 3.  
DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
FT NON\_TER 1  
FT NON\_TER 112  
SQ SEQUENCE 112 AA; 12638 MW; 4141BD7AD9924E7A CRC64;

Query Match 43.7%; Score 94; DB 4; Length 112;  
Best Local Similarity 41.2%; Pred. No. 1.3e-05;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECEIINHTNCNDVGYGPOCQ 34  
Db 47 CSSEPCKNNGTCVLDLITNRFNCCEPHGPFCEL 80

```
RESULT 19
P79941 PRELIMINARY; PRT; 642 AA.
AC P79941;
DT 01-MAY-1997 (TREMBLrel. 03, Created)
DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH LIGAND X-DELTA-2.
GN X-DELTA-2.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBL_TaxID=6355;
RN [1]
RP SEQUENCE FROM N.A.
RA Jen W.C., Weinstein D.A., Chitnis A.B., Kintner C.;
RT "The Notch ligand, X-Delta-2, mediates segmentation of the paraxial
RT mesoderm in Xenopus embryos.";
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
DR HSSP: P00740; 1EDM.
DR PROSITE: PS000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR SMART: SM00051; DSL; 1.
DR SMART: SM000179; EGF_CA; 3.
DR SMART: SM000001; EGF_like; 5.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_2.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 7.
DR PROSITE: PS01187; EGF_CA; 2.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SQ SEQUENCE 642 AA; 70667 MW; D7DC31EB9D92820C CRC64;

Query Match 43.3%; Score 93; DB 13; Length 642;
Best Local Similarity 41.2%; Pred. No. 0.00011;
Matches 14; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 1 CQWSCSGHGECEYIINHTNCNDVGYGPQCQL 34
DB 428 CASSPCANGSTCVDAVNSTYSCSTLGYGKDCFL 461

RESULT 20
O9UJV2 PRELIMINARY; PRT; 723 AA.
ID O9UJV2;
AC O9UJV2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DELTA-LIKE-1 PROTEIN.
GN DL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Han W., Ye Q., Moore M.A.S.;
RT "A Soluble Form of Human Delta-Like-1 Inhibits Differentiation of
RT Hematopoietic Progenitor Cells.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF196571; AAF05834.1; -.
```

```
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00051; DSL; 1.
DR SMART: SM000179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 7.
DR PROSITE: PS01187; EGF_CA; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SQ SEQUENCE 723 AA; 77985 MW; C7F41BD4265A5F7 CRC64;
```

```
Query Match 43.3%; Score 93; DB 4; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.00012;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQWSCSGHGECEYIINHTNCNDVGYGPQCQL 34
DB 332 CDSPCKNGSCSTDLNYSCTCPGFGKICEL 365
```

```
RESULT 21
O9NU41 PRELIMINARY; PRT; 723 AA.
ID O9NU41;
AC O9NU41;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE DJ894D12.3 (DELTA-LIKE 1 (MOUSE) HOMOLOG).
GN DL1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Homiidae; Homo.
OX NCBL_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Almeida J.;
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL078605; CAB89569.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01414; DSL; 1.
DR Pfam: PF00008; EGF; 8.
DR PRINTS: PR00010; EGFBL00D.
DR PRINTS: PR00011; EGF_LAMININ.
DR SMART: SM00051; DSL; 1.
DR SMART: SM00179; EGF_CA; 4.
DR SMART: SM00001; EGF_like; 4.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_3.
DR PROSITE: PS00022; EGF_1; UNKNOWN_8.
DR PROSITE: PS01186; EGF_2; 8.
DR PROSITE: PS01187; EGF_CA; 1.
KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
SQ SEQUENCE 723 AA; 78055 MW; 0948BF25D899D CRC64;
```

Query Match 43.3%; Score 93; DB 4; Length 723;  
 Best Local Similarity 38.2%; Pred. No. 0.00012;  
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPCOL 34  
 DB 332 CDPSPCKNGSCDLENSYTCPPGFGKICEL 365

## RESULT 22

O96NU3 PRELIMINARY; PRT; 2656 AA.  
 ID O96NU3;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE FIBROSURFIN PRECURSOR.  
 GN SURFIN2656.  
 OS Paracentrotus lividus (Common sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinidae;  
 OC Paracentrotus.  
 OX NCBI\_TaxID=7656;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=ADULT TEST;  
 RX MEDLINE=21264600; PubMed=11259425;  
 RA Cluvel C., Lechias C., Humbert F., Garrone R., Exposito J.Y.;  
 RT "Characterization of fibrosurin, an interfibrillar component of sea  
 urchin catch connective tissues.";  
 RL J. Biol. Chem. 276:18108-18114(2001).  
 DR EMBL: AJ291489; CAC20782.1; -.  
 DR HSSP: P00740; 1EDM.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR Pfam: PF00008; EGF; 17.  
 DR PRINTS: PR00010; EGFBLD.  
 DR SMART: SM00181; EGF; 17.  
 DR SMART: SM00179; EGF-Ca; 17.  
 DR SMART: SM00001; EGF-like; 5.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_1.  
 KW Signal.  
 FT SIGNAL.  
 SQ SEQUENCE 2656 AA; 276689 MW; 7CEC6A190886C229 CRC64;

Query Match 42.3%; Score 91; DB 5; Length 2656;  
 Best Local Similarity 39.4%; Pred. No. 0.00087;  
 Matches 13; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

OY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPCOL 33  
 DB 2556 CASGPCENAGDCIDDEVNSTCCTAGYBELYCO 2588

RESULT 23

O91902 PRELIMINARY; PRT; 721 AA.  
 ID O91902;  
 AC O91902;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE X-DELTA-1.  
 OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8335;  
 [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=95319507; PubMed=7596411;  
 RA Henrique D., Adam J., Wrat A., Chitnis A., Lewis J., Ish-Horowicz D.;  
 RT "Expression of a Delta homologue in prospective neurons in the  
 chick.";  
 RL Nature 375:787-790(1995).  
 [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=95319503; PubMed=7596407;  
 RA Chitnis A., Henrique D., Lewis J., Ish-Horowicz D., Kintner C.;  
 RT "Primary neurogenesis in Xenopus embryos regulated by a homologue of  
 the Drosophila neurogenic gene Delta.";  
 RL Nature 375:761-766(1995).  
 DR EMBL: L42229; AAC38017.1; -.  
 DR HSSP: P00740; 1EDM.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR001774; DSL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR Pfam: PF01414; DSL; 1.  
 DR Pfam: PF00008; EGF; 8.  
 DR PRINTS: PR00010; EGFBLD.  
 DR SMART: SM00051; DSL; 1.  
 DR SMART: SM00179; EGF-Ca; 4.  
 DR SMART: SM00001; EGF-like; 4.  
 DR PROSITE: PS00010; ASX\_HYDROXYL; UNKNOWN\_3.  
 DR PROSITE: PS00022; EGF-1; UNKNOWN\_8.  
 DR PROSITE: PS01186; EGF-2; 8.  
 DR PROSITE: PS01187; EGF-Ca; 2.  
 KW Calcium-binding; EGF-like domain; Glycoprotein; Repeat.  
 SQ SEQUENCE 721 AA; 79922 MW; 9EBDC65C439DD194 CRC64;

Query Match 41.9%; Score 90; DB 13; Length 721;  
 Best Local Similarity 38.2%; Pred. No. 0.00031;  
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCSGHCEVEIINHTCNCMDVGYGPCOL 34  
 DB 334 CDANPCKNGSCDLENSYTCPPGFGKNCLEL 367

RESULT 24

O16004 PRELIMINARY; PRT; 2531 AA.  
 ID O16004;  
 AC O16004;  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE NOTCH HOMOLOG.  
 OS Lytechinus variegatus (Sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidae; Euechinoidae; Echinacea; Temnopleurouida; Toxopneustidae;  
 OC Lytechinus.  
 OX NCBI\_TaxID=7654;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97454256; PubMed=9310331;  
 RA Sherwood D.R., McClay D.R.;  
 RT "Identification and localization of a sea urchin Notch homologue:  
 RT insights into vegetal plate regionalization and Notch receptor  
 regulation.";  
 RL Development 124:3363-3374(1997).  
 DR EMBL: AF000634; AAB82088.1; -.  
 DR HSSP: P01132; 1EGF.  
 DR InterPro: IPR002110; ANK.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF-Ca.  
 DR InterPro: IPR001438; EGF-II.  
 DR InterPro: IPR000800; Notch.

DR Pfam; PF00023; ank; 6.  
 DR Pfam; PF00008; EGF; 35.  
 DR Pfam; PF00066; notch; 3.  
 DR PRINTS; PR01415; ANKYRIN.  
 DR PRINTS; PR00010; EGFBLD.  
 DR PRINTS; PR01452; NOTCH.  
 DR SMART; SM00248; ANK; 5.  
 DR SMART; SM00179; EGF\_CA; 23.  
 DR SMART; SM00001; EGF\_like; 11.  
 DR SMART; SM00004; NL; 3.  
 DR PROSITE; PS50088; ANK\_REPEAT; 10.  
 DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 2.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; 21.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_33.  
 DR PROSITE; PS01186; EGF\_2; 25.  
 DR PROSITE; PS01187; EGF\_CA; 20.  
 DR ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;  
 KW Hydroxylation; Repeat.  
 SO SEQUENCE 2531 AA; 273982 MW; 5BF42BEC627CA303 CRC64;

Query Match 41.9%; Score 90; DB 5; Length 2531;  
 Best Local Similarity 38.2%; Pred. No. 0.0011;  
 Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPMSCGHECVELINNNHCNCDVGYGPOCOL 34  
 Db 555 CQSPRCENGCTCIDGVNQTCLCTGYGHRCEM 588

RESULT 25  
 ID Q9UA13 PRELIMINARY; PRT; 3767 AA.  
 AC Q9UA13; Q21340;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE TRANSMEMBRANE CELL ADHESION RECEPTOR MDA-3 PRECURSOR (K08E5.3 PROTEIN)  
 DE (FRAGMENT).  
 GN MDA-3 OR K08E5.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=N2;  
 RA Lu Z., Vogel B., Hedgecock E.;  
 RT "mda-3 mRNA Splicing Pattern Revealed";  
 RT Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Kershaw J.K.;  
 RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none.  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RL Science 282:2012-2018(1998).  
 DR EMBL; AF139060; AM29428.1; -  
 DR EMBL; Z30974; CA83226.2; -  
 DR EMBL; Z30423; CA83226.2; JOINED.  
 DR HSSP; P01130; ILDL.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR001762; Disintegrin.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001881; EGF\_CA.  
 DR InterPro; IPR002172; LDL\_recept\_A.  
 DR InterPro; IPR000082; SEA.  
 DR Pfam; PF00008; EGF; 33.  
 DR Pfam; PF00008; EGF; 33.

DR Pfam; PF00057; Idl\_recept\_a; 4.  
 DR Pfam; PF01390; SEA; 2.  
 DR Pfam; PF00092; vwa; 1.  
 DR PRINTS; PR00289; DISINTEGRIN.  
 DR PRINTS; PR00261; LDLRECEPTOR.  
 DR PRINTS; PR00453; VWFADOMAIN.  
 DR SMART; SM00181; EGF; 51.  
 DR SMART; SM00179; EGF\_CA; 1.  
 DR SMART; SM00001; EGF\_like; 45.  
 DR SMART; SM00192; LDLra; 4.  
 DR SMART; SM00200; SEA; 2.  
 DR SMART; SM00327; VWA; 1.  
 DR PROSITE; PS00010; ASX\_HYDROXYL; UNKNOWN\_32.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_11.  
 DR PROSITE; PS01186; EGF\_2; 6.  
 DR PROSITE; PS01187; EGF\_CA; 1.  
 DR PROSITE; PS50068; LDLRA\_2; 2.  
 DR PROSITE; PS50024; SEA; 4.  
 DR PROSITE; PS50234; VWEA; 2.  
 DR Calcium-binding; EGF-like domain; Glycoprotein; Receptor; Repeat;  
 KW Signal.  
 FT SIGNAL 1 24 POTENTIAL.  
 FT CRAIN 25 >3767 3.  
 FT NON\_TER 3767 3767  
 SO SEQUENCE 3767 AA; 417284 MW; 8DA3AE5EA50AE8E CRC64;

Query Match 41.9%; Score 90; DB 5; Length 3767;  
 Best Local Similarity 51.6%; Pred. No. 0.0017;  
 Matches 16; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

OY 6 CQSGHECVELINNNHCNCDVGYGPOCOL 34  
 Db 3380 CNAHGDCVHTATNNITCTDGTGPOCOV 3410

RESULT 26  
 ID Q9GPA5 PRELIMINARY; PRT; 2524 AA.  
 AC Q9GPA5;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PUTATIVE NOTCH RECEPTOR PROTEIN.  
 GN NOTCH.  
 OS Branchiostoma floridae (Florida lancelet) (Amphioxus).  
 OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
 OC Branchiostoma.  
 OX NCBI\_TaxID=7739;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=LARVA;  
 RA Holland L.Z., Burgdorf C., Holland N.D., Lehrach H., Tamme R.,  
 RA Abi-Rached L., Pontarotti P., Lardelli M.;  
 RT "Cloning and developmental expression of the amphioxus homologue of  
 RT Notch (Amphinotch): evolutionary conservation of multiple expression  
 RT domains in amphioxus and vertebrates";  
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; Y12539; CAC19873.1; -  
 DR HSSP; P00740; IEDM.  
 DR InterPro; IPR002110; ANK.  
 DR InterPro; IPR000152; Asx\_hydroxyl.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR000742; EGF\_2.  
 DR InterPro; IPR001881; EGF\_CA.  
 DR InterPro; IPR001438; EGF\_11.  
 DR InterPro; IPR002049; Laminin\_EGF.  
 DR InterPro; IPR000800; NOTCH.  
 DR Pfam; PF00008; EGF; 36.  
 DR Pfam; PF00066; notch; 3.  
 DR PRINTS; PR00010; EGFBLD.  
 DR PRINTS; PR00011; EGFLAMININ.

```
DR PRINTS: PRO1452; NOTCH.
DR SMART: SM00248; ANK: 6.
DR SMART: SM00181; EGF: 37.
DR SMART: SM00179; EGF_CA: 34.
DR SMART: SM00001; EGF_1like: 13.
DR SMART: SM00004; NL: 3.
DR PROSITE: PS50088; ANK_REPEAT: 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 1.
DR PROSITE: PS00010; ASX_HYDROXYL: 1.
DR ANK repeat: EGF-1like domain; Glycoprotein; Hydroxylation; Receptor;
KW Repeat.
SW SEQUENCE 2524 AA; 270969 MW; C2CA57E306D23EC9 CRC64;

Query Match 41.6%; Score 89.5; DB 5; Length 2524;
Best Local Similarity 45.5%; Pred. No. 0.0013;
Matches 15; Conservative 6; Mismatches 11; Indels 1; Gaps 1;

Oy 1 COPWSCSGHCEVEIINHTCNCVDVGYGPOCQ 33
Db 559 COSNCCQ-HGTCVGVASTCCEPGYNGPLCE 590

RESULT 27
O06008 PRELIMINARY; PRT; 1203 AA.
AC O06008.
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH B PROTEIN (FRAGMENT).
GN NOTCH2 OR NOTCH B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=FL (CBA X C57BL); TISSUE=EMBRO;
RX MEDLINE=93178563; PubMed=8440332;
RA Lardelli M., Lendahl U.;
RT "Notch A and Notch B - two mouse Notch homologues coexpressed in a
RT wide variety of tissues."
RL Exp. Cell Res. 204:364-372(1993).
DR EMBL: X68279; CAA48340.1; -.
DR HSSP: P16109; IFSB.
DR MGD: MGI:97364; Notch2.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF_1like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR000083; fibronectin_type_1.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00008; EGF: 27.
DR Pfam: PF00066; notch: 1.
DR PRINTS: PRO0010; EGFBL00D.
DR PRINTS: PRO0011; EGF_LAMININ.
DR PRINTS: PRO0012; FNTYPEI.
DR SMART: SM00179; EGF_CA: 19.
DR SMART: SM00001; EGF_1like: 7.
DR SMART: SM00004; NL: 2.
DR PROSITE: PS00010; ASX_HYDROXYL: 19.
DR PROSITE: PS00022; EGF_1; UNKNOWN_27.
DR PROSITE: PS01186; EGF_2; 22.
DR PROSITE: PS01187; EGF_CA: 19.
DR Calcium-binding: EGF-1like domain; Glycoprotein; Hydroxylation; Repeat.
KW NON_TER 1
FT SEQUENCE 1203 AA; 128982 MW; 65040843799BCA37 CRC64;

Query Match 41.4%; Score 89; DB 11; Length 1203;
```

```
Best Local Similarity 41.2%; Pred. No. 0.00072;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHCEVEIINHTCNCVDVGYGPOCQ 34
Db 181 COSNPCVNNQCVDKYNRFQCLCPGFTGPVCQI 214

RESULT 28
O09UPL3 PRELIMINARY; PRT; 2281 AA.
AC O09UPL3.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH 3 (FRAGMENT).
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Lamerdin J.E., McCreedy P.M., Skowronski E., Adamson A.W.,
RA Burkhardt-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA Phan H., Velasco N., Ganes J., Dangnan L., Poundstone P.,
RA Christensen M., Georges A., Avila J., Liu S., Atlix C., Andreise T.,
RA Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA Bruce R., Thomas P., Quan G., Krommiller B., Arellano A.,
RA Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.S.,
RA Carrano A.V.;
RT "Sequence analysis of an ~1.5 Mb olfactory receptor (OLFR) cluster in
RT 19p13.1."
RL Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AC004663; AAC15789.1; -.
DR HSSP: P00740; 1EDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001762; Disintegrin.
DR InterPro: IPR000561; EGF_1like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_CA.
DR InterPro: IPR001438; EGF_11.
DR InterPro: IPR000083; fibronectin_type_1.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank: 6.
DR Pfam: PF00008; EGF: 34.
DR Pfam: PF00066; notch: 3.
DR PRINTS: PRO0289; DISINTEGRIN.
DR PRINTS: PRO0010; EGFBL00D.
DR PRINTS: PRO0011; EGF_LAMININ.
DR PRINTS: PRO0012; FNTYPEI.
DR SMART: SM00248; ANK: 5.
DR SMART: SM00179; EGF_CA: 19.
DR SMART: SM00001; EGF_1like: 14.
DR SMART: SM00004; NL: 3.
DR PROSITE: PS50088; ANK_REPEAT: 8.
DR PROSITE: PS50297; ANK_REPEAT_REGION: 2.
DR PROSITE: PS00010; ASX_HYDROXYL: 18.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 25.
DR PROSITE: PS01187; EGF_CA: 16.
DR ANK repeat: Calcium-binding; EGF-1like domain; Glycoprotein;
KW Hydroxylation; Repeat.
KW NON_TER 1
FT SEQUENCE 2281 AA; 239472 MW; DE23B2F129709D75 CRC64;

Query Match 41.4%; Score 89; DB 4; Length 2281;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
```

```

QY 1 CQWSCSGHGECEVEIINNHTCNCADVGYGPQCQ 33
DB 888 CSPSSCFNGTCVCDGVNSFSCICRPGYTGACQC 920

RESULT 29
Q9R172 PRELIMINARY; PRT; 2319 AA.
AC Q9R172;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NOTCH 3 PROTEIN.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Haitianians T., Boulter J., Weimaster G., Schanen N.C.;
RT "Rattus norvegicus mRNA for Notch 3."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF164486; AAD4653.2; -
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 33.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PRO0010; EGFBLD.
DR PRINTS; PRO0011; EGFAMININ.
DR PRINTS; PRO1452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 20.
DR SMART; SM00001; EGF-like; 12.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 1.
DR ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
KW SEQUENCE 2319 AA; 244298 MW; 243BCA02D7C3283D CRC64;
SQ

Query Match 41.4%; Score 89; DB 11; Length 2319;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
QY 1 CQWSCSGHGECEVEIINNHTCNCADVGYGPQCQ 33
DB 930 CSPSSCFNGTCVCDGVNSFSCICRPGYTGACQC 962

RESULT 30
Q9Y6L8 PRELIMINARY; PRT; 2321 AA.
AC Q9Y6L8;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NOTCH3.
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.

```

```

RA Gunel M., Artavanis-Tsakonas S.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF058900; AAC14346.1; -
DR EMBL; AF058881; AAC14346.1; JOINED.
DR EMBL; AF058882; AAC14346.1; JOINED.
DR EMBL; AF058883; AAC14346.1; JOINED.
DR EMBL; AF058884; AAC14346.1; JOINED.
DR EMBL; AF058885; AAC14346.1; JOINED.
DR EMBL; AF058886; AAC14346.1; JOINED.
DR EMBL; AF058887; AAC14346.1; JOINED.
DR EMBL; AF058888; AAC14346.1; JOINED.
DR EMBL; AF058889; AAC14346.1; JOINED.
DR EMBL; AF058890; AAC14346.1; JOINED.
DR EMBL; AF058891; AAC14346.1; JOINED.
DR EMBL; AF058892; AAC14346.1; JOINED.
DR EMBL; AF058893; AAC14346.1; JOINED.
DR EMBL; AF058894; AAC14346.1; JOINED.
DR EMBL; AF058895; AAC14346.1; JOINED.
DR EMBL; AF058896; AAC14346.1; JOINED.
DR EMBL; AF058897; AAC14346.1; JOINED.
DR EMBL; AF058898; AAC14346.1; JOINED.
DR EMBL; AF058899; AAC14346.1; JOINED.
DR HSSP; P00740; 1EDM.
DR InterPro; IPR002110; ANK.
DR InterPro; IPR00152; Asx_hydroxyl.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR000742; EGF_2.
DR InterPro; IPR001881; EGF_Ca.
DR InterPro; IPR001438; EGF_II.
DR InterPro; IPR002049; Laminin_EGF.
DR InterPro; IPR000800; Notch.
DR Pfam; PF00023; ank; 6.
DR Pfam; PF00008; EGF; 34.
DR Pfam; PF00066; notch; 3.
DR PRINTS; PRO0010; EGFBLD.
DR PRINTS; PRO0011; EGFAMININ.
DR PRINTS; PRO1452; NOTCH.
DR SMART; SM00248; ANK; 5.
DR SMART; SM00179; EGF_CA; 19.
DR SMART; SM00001; EGF-like; 14.
DR PROSITE; PS50088; ANK_REPEAT; 4.
DR PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE; PS00010; ASX_HYDROXYL; 18.
DR PROSITE; PS00022; EGF_1; UNKNOWN_33.
DR PROSITE; PS01186; EGF_2; 25.
DR PROSITE; PS01187; EGF_CA; 16.
KW ANK repeat; Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2321 AA; 243628 MW; 3E70EC12A59CD638 CRC64;

Query Match 41.4%; Score 89; DB 4; Length 2321;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
QY 1 CQWSCSGHGECEVEIINNHTCNCADVGYGPQCQ 33
DB 928 CSPSSCFNGTCVCDGVNSFSCICRPGYTGACQC 960

RESULT 31
Q9UM47 PRELIMINARY; PRT; 2321 AA.
AC Q9UM47;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE NOTCH3.
GN NOTCH3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

```



```
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97032728; PubMed=8678478;
RA Joutel A., Corpechot C., Ducros A., Vahedi K., Chabriat H., Mouton P.,
RA Almqvist S., Domenga V., Cecillon M., Marechal E., Maciazek J.,
RA Vaysiere C., Cruaud C., Cabanis E.A., Ruchoux M.M., Weissbach J.,
RA Bach J.F., Bousser M.G., Tournier-Lasserre E.;
RT "Notch3 mutations in CADASIL, a hereditary adult-onset condition
RT causing stroke and dementia.";
RL Nature 383:707-710(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98049753; PubMed=9388399;
RA Joutel A., Vahedi K., Corpechot C., Trosch A., Chabriat H.,
RA Vaysiere C., Cruaud C., Maciazek J., Weissbach J., Bousser M.G.,
RA Bach J.F., Tournier-Lasserre E.;
RT "Strong clustering and stereotyped nature of Notch3 mutations in
RT CADASIL patients.";
RL Lancet 350:1511-1515(1997).
DR EMBL: U97669; AAB91371.1; -.
DR HSSP: P00740; IEDM.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR002049; Laminin_EGF.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 34.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFBLDOD.
DR PRINTS: PR00011; EGFBLAMIN.
DR PRINTS: PR01452; NOTCH.
DR SMART: SM00248; ANK; 5.
DR SMART: SM00179; EGF_CA; 19.
DR SMART: SM00001; EGF_Like; 14.
DR SMART: SM00004; NL; 3.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 18.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 25.
DR PROSITE: PS01187; EGF_CA; 16.
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2321 AA; 243656 MW; AFAEF2ECAB7C836 CRC64;
```

```
Query Match 41.4%; Score 89; DB 4; Length 2321;
Best Local Similarity 45.5%; Pred. No. 0.0014;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;
```

```
OY 1 CQWSCSGHGEVETIINHTCNDVGYGPQCQ 33
DB 928 CSPSSCFNGTCTGVGNSFSCRCRGYGAHCQ 960
```

```
RESULT 32
035516 PRELIMINARY: PRT: 2470 AA.
ID 035516
AC 035516
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CELL SURFACE PROTEIN.
GN NOTCH2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

```
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B/6; TISSUE=THYMUS;
RX MEDLINE=93178563; PubMed=8440332;
RA Lardelli M., Landahl U.;
RT "Notch A and Notch B - two mouse Notch homologues coexpressed in a
RT wide variety of tissues.";
RL Exp. Cell Res. 204:364-372(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57B/6; TISSUE=THYMUS;
RA Hamada Y., Higuchi M., Tsujimoto Y.;
RT "Complete amino acid sequence and multiform transcripts encoded by a
RT single copy of mouse Notch2 gene.";
RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
DR EMBL: D32210; BAA22094.1; -.
DR HSSP: P16109; IFSB.
DR MGD: MGI:97364; Notch2.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 35.
DR Pfam: PF00066; notch; 2.
DR PRINTS: PR00010; EGFBLDOD.
DR PRINTS: PR01452; NOTCH.
DR SMART: SM00248; ANK; 4.
DR SMART: SM00179; EGF_CA; 22.
DR SMART: SM00001; EGF_Like; 12.
DR SMART: SM00004; NL; 3.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 22.
DR PROSITE: PS00022; EGF_1; UNKNOWN_33.
DR PROSITE: PS01186; EGF_2; 27.
DR PROSITE: PS01187; EGF_CA; 22.
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;
KW Hydroxylation; Repeat.
SQ SEQUENCE 2470 AA; 265325 MW; B5A31B35242716D CRC64;
```

```
Query Match 41.4%; Score 89; DB 11; Length 2470;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;
```

```
OY 1 CQWSCSGHGEVETIINHTCNDVGYGPQCQ 34
DB 496 COSNPCVNGGCVCKVYKVRFCCLCPGFTGPVCOI 529
```

```
RESULT 33
004721 PRELIMINARY: PRT: 2471 AA.
AC 004721
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NOTCH 2 PRECURSOR.
GN N2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE OF 1810-2447 FROM N.A.
RC TISSUE=BRIN;
RX MEDLINE=9365135; PubMed=1303260;
RA Stifani S., Bismuth C.M., Redhead N.J., Hill R.E.,
RA Artavanis-Tsakonas S.;
```

```

RT      "Human homologs of a Drosophila enhancer of split gene product define
RT      a novel family of nuclear proteins."
RL      Nat. Genet. 2:119-127(1992).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE-BRAIN:
RA      Blaumuller C.M., Mann R.S.;
RT      "Complete Human Notch 2 (hN2) cDNA sequence.";
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF308601; AAA36377.2; -.
DR      HSSP; P00740; IEDM.
DR      InterPro; IPR002110; ANK.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR000742; EGF_2.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR001438; EGF_II.
DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR000800; Notch.
DR      Pfam; PF00023; ank; 12.
DR      Pfam; PF00008; EGF; 35.
DR      Pfam; PF00066; notch; 2.
DR      PRINTS; PR00010; EGFBL00D.
DR      PRINTS; PR00011; EGFLAMININ.
DR      SMART; SM00248; ANK; 4.
DR      SMART; SM00179; EGF_CA; 35.
DR      SMART; SM00001; EGF_Like; 12.
DR      SMART; SM00004; NL; 2.
DR      PROSITE; PS50088; ANK_REPEAT; 4.
DR      PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR      PROSITE; PS00010; ASX_HYDROXYL; 1.
DR      ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat;
KW      Signal.
FT      SIGNAL.
SQ      SEQUENCE 2471 AA; 265401 MW; DB4FD003CCE5AF5 CRC64;

Query Match          41.4%; Score 89; DB 4; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY      1 CQWSCSGHGEVCYEIINNHTCNCVDVGYGPQCL 34
Db      498 CQSNPCVNGCVDKVNRFQCLCPGFTGPVCOI 531

RESULT 34
ID      Q9H240; PRELIMINARY; PRT; 2471 AA.
AC      Q9H240;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      NOTCH2 PROTEIN.
GN      NOTCH2.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON      NCBI_TaxID=9606;
RX      Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF315356; AAG37073.1; -.
DR      HSSP; P00740; IEDM.
DR      InterPro; IPR002110; ANK.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR000742; EGF_2.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR001438; EGF_II.
DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR000800; Notch.
DR      Pfam; PF00023; ank; 12.
DR      Pfam; PF00008; EGF; 35.
DR      Pfam; PF00066; notch; 2.
DR      PRINTS; PR00010; EGFBL00D.
DR      PRINTS; PR00011; EGFLAMININ.
DR      SMART; SM00248; ANK; 4.
DR      SMART; SM00179; EGF_CA; 35.
DR      SMART; SM00001; EGF_Like; 12.
DR      SMART; SM00004; NL; 2.
DR      PROSITE; PS50088; ANK_REPEAT; 4.
DR      PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR      PROSITE; PS00010; ASX_HYDROXYL; 1.
DR      ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ      SEQUENCE 2471 AA; 265429 MW; 7DA7F8F44589B1C CRC64;

```

```

DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR000800; Notch.
DR      Pfam; PF00008; EGF; 32.
DR      Pfam; PF00066; notch; 2.
DR      PRINTS; PR00010; EGFBL00D.
DR      PRINTS; PR00011; EGFLAMININ.
DR      PRINTS; PR01452; NOTCH.
DR      SMART; SM00248; ANK; 6.
DR      SMART; SM00181; EGF; 37.
DR      SMART; SM00179; EGF_CA; 35.
DR      SMART; SM00001; EGF_Like; 12.
DR      SMART; SM00004; NL; 2.
DR      PROSITE; PS50088; ANK_REPEAT; 4.
DR      PROSITE; PS50297; ANK_REPEAT_REGION; 1.
DR      PROSITE; PS00010; ASX_HYDROXYL; 1.
DR      ANK repeat; EGF-like domain; Glycoprotein; Hydroxylation; Repeat.
SQ      SEQUENCE 2471 AA; 265429 MW; 7DA7F8F44589B1C CRC64;

Query Match          41.4%; Score 89; DB 4; Length 2471;
Best Local Similarity 41.2%; Pred. No. 0.0015;
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY      1 CQWSCSGHGEVCYEIINNHTCNCVDVGYGPQCL 34
Db      498 CQSNPCVNGCVDKVNRFQCLCPGFTGPVCOI 531

RESULT 35
ID      Q9QW30; PRELIMINARY; PRT; 2471 AA.
AC      Q9QW30;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      NOTCH2 PROTEIN.
GN      Rattus sp., and
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON      NCBI_TaxID=10118, 10116;
RX      Submitted (APR-1993) to the EMBL/GenBank/DBJ databases.
DR      EMBL; M93661; AAK13558.1; -.
DR      HSSP; P00743; ICCF.
DR      InterPro; IPR002110; ANK.
DR      InterPro; IPR000152; Asx_hydroxyl.
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR000742; EGF_2.
DR      InterPro; IPR001881; EGF_Ca.
DR      InterPro; IPR001438; EGF_II.
DR      InterPro; IPR002049; Laminin_EGF.
DR      InterPro; IPR000800; Notch.
DR      Pfam; PF00023; ank; 6.
DR      Pfam; PF00008; EGF; 35.
DR      Pfam; PF00066; notch; 2.
DR      PRINTS; PR00010; EGFBL00D.
DR      PRINTS; PR00011; EGFLAMININ.

```

DR PRINTS: PRO1452; NOTCH.  
DR SMART: SM00248; ANK: 4.  
DR SMART: SM00179; EGF\_CA: 22.  
DR SMART: SM00001; EGF\_Like: 10.  
DR SMART: SM00004; NL: 2.  
DR PROSITE: PS50088; ANK\_REPEAT: 4.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION: 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL: 22.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_34.  
DR PROSITE: PS01186; EGF\_2; 26.  
DR PROSITE: PS01187; EGF\_CA: 22.  
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;  
KW Hydroxylation; Repeat.  
SQ SEQUENCE 2471 AA; 265367 MW; 7D5C8E18DDE95F8 CRC64;

Query Match 41.4%; Score 89; DB 11; Length 2471;  
Best Local Similarity 41.2%; Pred. No. 0.0015;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

OY 1 CPWSCSGHGCVEIINHTCNCDDVGYGPOCQ 34  
DB 498 CQSNPCVNNGCVDKVNRFQCCLPPTGTCVCOI 531

RESULT 36  
025253 PRELIMINARY; PRT; 2653 AA.  
AC 025253;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NOTCH HOMOLOG SCALLOPED WINGS (SCL).  
GN SCL.  
OS Lucilia cuprina (Greenbottle fly) (Australian sheep blowfly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Cecidomyiidae; Calliphoridae; Lucilla.  
OC NCBI\_TaxID=7375;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SS SEEKING;  
RX MEDLINE=96400928; Pubmed=8807304;  
RA Davies A.G., Game A.Y., Chen Z., Williams T.J., Goodall S., Yen J.L.,  
RA McKenzie J.A., Batterham P.;  
RT "Scalloped wings is the Lucilla cuprina Notch homologue and a  
RT candidate for the modifier of fitness and asymmetry of diazino  
RT resistance.";  
RL Genetics 143:1321-1337(1996).  
RN [2]  
RP SEQUENCE OF 39-265 FROM N.A.  
RC STRAIN=SS SEEKING;  
RA Chen Z., Newsome T., McKenzie J.A., Batterham P.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 39-265 FROM N.A.  
RC STRAIN=SS SEEKING;  
RA Chen Z., McKenzie J.A., Batterham P.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U58977; AAC36151.1; -  
DR EMBL: AF032672; AAC36152.1; -  
DR EMBL: AF032670; AAC36152.1; JOINED.  
DR EMBL: AF032671; AAC36152.1; JOINED.  
DR EMBL: AF032673; AAC36153.1; -  
DR HSSP: P00740; LEDM.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF\_2.  
DR InterPro: IPR001881; EGF\_CA.  
DR InterPro: IPR001438; EGF\_11.  
DR InterPro: IPR000800; Notch.  
DR Pfam: PF00023; ank; 6.

DR Pfam: PF00008; EGF; 36.  
DR Pfam: PF00066; notch; 3.  
DR PRINTS: PRO0010; EGFBLD.  
DR PRINTS: PRO1452; NOTCH.  
DR SMART: SM00248; ANK: 4.  
DR SMART: SM00179; EGF\_CA: 24.  
DR SMART: SM00001; EGF\_Like; 11.  
DR SMART: SM00004; NL: 2.  
DR PROSITE: PS50088; ANK\_REPEAT: 5.  
DR PROSITE: PS50297; ANK\_REPEAT\_REGION: 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL: 22.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_34.  
DR PROSITE: PS01186; EGF\_2; 28.  
DR PROSITE: PS01187; EGF\_CA: 21.  
DR ANK repeat: Calcium-binding; EGF-like domain; Glycoprotein;  
KW Hydroxylation; Repeat.  
SQ SEQUENCE 2653 AA; 285928 MW; 6AF2A058FEEC6329 CRC64;

Query Match 41.4%; Score 89; DB 5; Length 2653;  
Best Local Similarity 45.2%; Pred. No. 0.0016;  
Matches 14; Conservative 7; Mismatches 10; Indels 0; Gaps 0;

OY 3 PWSGSGHGCVEIINHTCNCDDVGYGPOCQ 33.  
DB 1286 PCSNAGTLDVCVQVNNHNCCKPGYGRHCE 1316

RESULT 37  
093574 PRELIMINARY; PRT; 3209 AA.  
ID 093574  
AC 093574;  
DT 01-NOV-1998 (TREMBLrel. 08, Created)  
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE EXTRACELLULAR REELIN (FRAGMENT).  
GN RELN.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OC NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Bernier B., Goffinet A.M.;  
RT "Comparative study of reelin in vertebrates.";  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF090441; AAC35559.1; -  
DR InterPro: IPR002860; BNR.  
DR InterPro: IPR000561; EGF-like.  
DR Pfam: PF02012; BNR; 15.  
DR Pfam: PF00008; EGF; 4.  
DR SMART: SM00181; EGF; 5.  
DR PROSITE: PS00022; EGF\_1; UNKNOWN\_7.  
DR PROSITE: PS01186; EGF\_2; 7.  
KW EGF-like domain; Glycoprotein.  
FT NON\_TER 1  
SQ SEQUENCE 3209 AA; 361288 MW; 81A7B6676BCAA3D1 CRC64;

Query Match 41.2%; Score 88.5; DB 13; Length 3209;  
Best Local Similarity 51.7%; Pred. No. 0.0023;  
Matches 15; Conservative 3; Mismatches 8; Indels 3; Gaps 1;

OY 6 CSGHGCVEIINHTCNCDDVGYGPOCQ 34  
DB 1885 CNGHSC---INGTKICIDPGYGPCTKI 1910

RESULT 38  
025058 PRELIMINARY; PRT; 529 AA.  
AC 025058;

```

DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE FIBROPELIN IA (FRAGMENT).
OS Helicoidaris erythrogramma (Sea urchin).
OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;
OC Echinoidae; Euechinoidae; Echinacea; Echinoida; Echinometridae;
OC Helicoidaris.
OX NCBI_TaxID=7634;
RN [1]
RP SEQUENCE FROM N.A.
RA Bisgrove B.W.;
RT "Evolution of the fibropellin gene family and pattern of fibropellin
RT gene expression in sea urchin phylogeny.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: L33861; AAA2995.1; -
DR HSP: P08709; 1BF9.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000088; Avidin.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF_2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000083; fibronectin_type_1.
DR InterPro: IPR002049; Laminin_EGF.
DR Pfam: PF01382; Avidin; 1.
DR PRINTS: PR00010; EGFBLD.
DR PRINTS: PR00011; EGFFLAMIN.
DR PRINTS: PR00012; FNTYPEI.
DR PRODOM: PD016055; Avidin; 1.
DR SMART: SM00179; EGF_CA; 9.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_8.
DR PROSITE: PS00577; AVIDIN; 1.
DR PROSITE: PS00022; EGF_1; UNKNOWN_11.
DR PROSITE: PS01186; EGF_2; 10.
DR PROSITE: PS01187; EGF_CA; 7.
DR Calcium-binding; EGF-like domain; Glycoprotein; Repeat.
FT NON_TER 1
SO SQUONCE 529 AA; 55543 MW; D4AE958FCF9ACB5A CRC64;

Query Match 40.9%; Score 88; DB 5; Length 529;
Best Local Similarity 36.4%; Pred. No. 0.00043;
Matches 12; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWCSGHEGVEIINHTCNCDCVGYGPOCQ 33
DB 215 CASMPLNGGCIEMWNGYTCGCAAGFTGVLC 247

RESULT 39
Q9VOI2 PRELIMINARY; PRT; 562 AA.
AC Q9VOI2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE CG2991 PROTEIN.
GN CG2991.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RA MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S.E., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

```

```

RA Brandon R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agdayani A., An H.-J., Andrews-Plankkosh C., Baldwin D.,
RA Baller R.M., Basu A., Bakendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Borchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahle C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Gary N.S., Gelbart W.M., Glaeser K.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Mishina N.V., Modarity C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Sytzkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach C.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL EMBL: AE003581; AAF51189.1; -
DR EMBL: FB90031474; CG2991.
DR InterPro: IPR000561; EGF-like.
DR PROSITE: PS00022; EGF_1; UNKNOWN_1.
SO SEQUENCE 562 AA; 63026 MW; B377A00FA17761DA CRC64;

Query Match 40.9%; Score 88; DB 5; Length 562;
Best Local Similarity 50.0%; Pred. No. 0.00045;
Matches 15; Conservative 5; Mismatches 6; Indels 4; Gaps 1;

QY 5 SCSGHGECEIINHTCNCDCVGYGPOCQ 34
DB 50 SCSGRDCV---NNTCVCIDIRYAGNECDI 75

RESULT 40
Q9UOE2 PRELIMINARY; PRT; 585 AA.
AC Q9UOE2;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIGNAL RECEPTOR PROTEIN (FRAGMENT).
GN NOTCH.
OS Tribolium castaneum (Red flour beetle).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Polyphaga;
OC Cucujiformia; Tenebrionidae; Tribolium.
OX NCBI_TaxID=7070;
RN [1]
RP SEQUENCE FROM N.A.
RA Tautz D., Lardelli M., Westin J., Tamme R.;
RT "Embryonic expression of Tribolium Notch.";
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ005083; CAB65469.1; -
DR HSP: P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.

```



**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:41 : Search time 18.65 Seconds  
(without alignments)  
70.588 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193  
Perfect score: 215  
Sequence: 1 CQPMSCSGHGEVCEIINNHTCNCNDVGYRQCQL 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	205	95.3	372	1	LEM1_HUMAN
2	205	95.3	372	1	LEM1_HUMAN
3	205	95.3	372	1	LEM1_MACMU
4	205	95.3	372	1	LEM1_PANTR
5	205	95.3	372	1	LEM1_PAPHA
6	191	88.8	370	1	LEM1_PONRY
7	169	78.6	372	1	LEM1_BOVIN
8	159	74.0	372	1	LEM1_BOVIN
9	131	60.9	611	1	LEM1_RAT
10	130	60.5	610	1	LEM2_CANFA
11	127	59.1	549	1	LEM2_HUMAN
12	127	59.1	549	1	LEM2_RAT
13	127	59.1	646	1	LEM3_BOVIN
14	126	58.6	769	1	LEM3_SHEEP
15	125	58.1	612	1	LEM2_MOUSE
16	125	58.1	484	1	LEM2_PIG
17	124	57.7	768	1	LEM3_HUMAN
18	118	54.9	485	1	LEM3_MOUSE
19	116	54.0	551	1	LEM2_BOVIN
20	109	50.7	768	1	LEM2_RABIT
21	95	44.2	603	1	LEM3_RAT
22	93	43.3	723	1	FA12_CAVPO
23	93	43.3	723	1	DILL1_HUMAN
24	90	41.9	2524	1	NOTC_XENLA
25	87	40.5	3051	1	YMX3_CAEEL
26	87	40.5	461	1	YMX3_CAEEL
27	87	40.5	1064	1	PRTC_MOUSE
28	86	40.0	2139	1	DL1L_RAT
29	86	40.0	2703	1	FBP1_STRPU
30	85.5	39.8	4703	1	CRB_DROME
31	84	39.1	2201	1	NOTC_DROME
32	84	39.1	2318	1	FP2_MITGA
33	83	38.6	461	1	TENA_HUMAN
					NTC3_MOUSE
					PRTC_RAT
					P31394_rattus norv

## ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	P14151	P15023			
DT	01-JAN-1990	(Rel. 13, Created)			
DT	01-FEB-1991	(Rel. 17, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SEL1 OR LYMA1 OR LNHR.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89315837; PubMed=2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary				
RL	conservation at tandem cell interaction domains.";				
RN	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Tonsil;				
RX	MEDLINE=89310350; PubMed=2473156;				
RA	Tedder T.F., Isaacs C.M., Ernst T.J., Demetri G.D., Adler D.A.,				
RT	Disteche C.M.;				
RT	"Isolation and chromosomal localization of cDNAs encoding a novel				
RT	human lymphocyte cell surface molecule, LAM-1. Homology with the				
RT	mouse lymphocyte homing receptor and other human adhesion proteins.";				
RN	J. Exp. Med. 170:123-133(1989).				
RN	[3]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90044046; PubMed=2509393;				
RA	Camerlin D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing				
RT	receptor.";				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89308881; PubMed=2663882;				
RA	Bower B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph				
RT	node homing receptor.";				
RL	J. Cell Biol. 103:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90243637; PubMed=1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RT	"Structure of the gene encoding the human leukocyte adhesion				
RT	molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";				
RN	J. Biol. Chem. 265:7760-7767(1990).				
RN	[6]				
RP	3D-STRUCTURE MODELING.				

P49013 strongyloce  
P82279 homo sapien  
O15543 drosophila  
P46530 brachydantio  
P00741 homo sapien  
Q27591 drosophila  
P46531 homo sapien  
O01705 mus musculu  
O07008 rattus norv  
O61483 mus musculu  
P01132 mus musculu  
P31695 mus musculu

RA MEDLINE-96074584; PubMed-7488174;  
 RT Bajorath J., Anuflo A.;  
 RT "A template for generation and comparison of three-dimensional  
 RT selectin models."  
 RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995)  
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE  
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
 CC VENULES IN PERIPHERAL LYMPH NODES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- DATABASE: MAME-PROW; NOTE=CD guide CD62L entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62l.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M25280; AAC63053.1; -;  
 DR EMBL: X16150; CAB43475.1; -;  
 DR EMBL: X17519; CAB43536.1; -;  
 DR EMBL: X17519; CAB43537.1; ALT\_SEQ.  
 DR EMBL: X16070; CAB34203.1; ALT\_SEQ.  
 DR EMBL: M32414; AAB60700.1; -;  
 DR EMBL: M32406; AAB60700.1; JOINED.  
 DR EMBL: M32407; AAB60700.1; JOINED.  
 DR EMBL: M32408; AAB60700.1; JOINED.  
 DR EMBL: M32409; AAB60700.1; JOINED.  
 DR EMBL: M32410; AAB60700.1; JOINED.  
 DR EMBL: M32411; AAB60700.1; JOINED.  
 DR EMBL: M32412; AAB60700.1; JOINED.  
 DR EMBL: M32413; AAB60700.1; JOINED.  
 DR PIR: A33912; A33912.  
 DR PIR: A34015; A34015.  
 DR PIR: S06798; S06798.  
 DR PIR: J10104; J10104.  
 DR PDB: 1KJB; 03-APR-96.  
 DR GLYCOSITEDB: P14151; -.  
 DR MTM: 153240; -;  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sushi; 2.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 2.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS00186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE\_LECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sushi; Repeat; 3D-structure.  
 FT SIGNAL 1 28  
 FT PROPEP 29 38  
 FT CHAIN 39 372  
 FT DOMAIN 39 332  
 FT TRANSMEM 333 355  
 FT DOMAIN 356 372  
 FT DOMAIN 355 155  
 FT DOMAIN 156 192  
 FT DOMAIN 196 255  
 FT DOMAIN 258 317  
 FT DISULFID 57 155  
 L-SELECTIN.  
 EXTRACELLULAR (POTENTIAL).  
 POTENTIAL.  
 CYTOPLASMIC (POTENTIAL).  
 C-TYPE LECTIN (SHORT FORM).  
 EGF-LIKE.  
 SUSHI 1.  
 SUSHI 2.  
 BY SIMILARITY.

FT	DISULFID	128	147	BY SIMILARITY.
FT	DISULFID	160	171	BY SIMILARITY.
FT	DISULFID	165	180	BY SIMILARITY.
FT	DISULFID	182	191	BY SIMILARITY.
FT	DISULFID	197	241	BY SIMILARITY.
FT	DISULFID	227	254	BY SIMILARITY.
FT	DISULFID	259	303	BY SIMILARITY.
FT	DISULFID	289	316	BY SIMILARITY.
FT	CARBOHYD	60	60	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	104	104	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	177	177	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	232	232	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	246	246	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	271	271	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	311	311	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	37	37	D -> Y (IN REF. 4).
FT	CONFLICT	178	178	Y -> H (IN REF. 4).
FT	CONFLICT	193	193	F -> L (IN REF. 1 AND 4).
FT	CONFLICT	213	213	P -> S (IN REF. 3).
FT	CONFLICT	214	214	L -> F (IN REF. 4).
FT	CONFLICT	218	220	SFS -> NFN (IN REF. 2).
FT	CONFLICT	242	242	G -> E (IN REF. 2).
SO	SEQUENCE	372 AA;	42187 MW;	6EA9918ECA2D3643 CRC64;

Query Match 95.3%; Score 205; DB 1; Length 372;  
 Best Local Similarity 97.0%; Pred. No. 1, 8e-18;  
 Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCGHECEVEIINNHTCNDVGYGPOCQ 33  
 Db 160 CQPMSCGHECEVEIINNHTCNDVGYGPOCQ 192

RESULT 2  
 LEAF\_MACMU STANDARD; PRT; 372 AA.  
 ID LEAF\_MACMU  
 AC Q95198;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).  
 DE SEIL.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecoidea; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,  
 RA Tsurushita N.;  
 RU Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE  
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
 CC VENULES IN PERIPHERAL LYMPH NODES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U73730; AAB18246.1; -;



```
DR HSP: P14151; IKJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE-LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
FT DISULFID 160 171
FT DISULFID 165 180
FT DISULFID 182 191
FT DISULFID 197 241
FT DISULFID 227 254
FT DISULFID 259 303
FT DISULFID 289 316
FT CARBOHD 60 60
FT CARBOHD 104 104
FT CARBOHD 177 177
FT CARBOHD 226 226
FT CARBOHD 232 232
FT CARBOHD 246 246
FT CARBOHD 271 271
SQ SEQUENCE 372 AA; 42109 MW; 08BD7DD5AC349D6D CRC64;

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1.8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEIINHTNCMDVGYGPOCQ 33
DB 160 CQPMSCSGHGECEIINHTNCMDVGYGPOCQ 192

RESULT 3
LEMI_PANTR STANDARD; PRT; 372 AA.
AC 095237;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (LECAM1) (CD62L).
GN SELL.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
```

```
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Tsurushita N.;
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73728; AB18248.1; -.
DR HSP: P14151; IKJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF; 1.
DR Pfam: PF00059; lectin_c; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS00615; C-TYPE-LECTIN_1; 1.
DR PROSITE: PS00615; C-TYPE-LECTIN_2; 1.
DR PROSITE: PS00615; C-TYPE-LECTIN_2; 1.
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
FT DISULFID 160 171
FT DISULFID 165 180
FT DISULFID 182 191
FT DISULFID 197 241
FT DISULFID 227 254
FT DISULFID 259 303
FT DISULFID 289 316
FT CARBOHD 60 60
FT CARBOHD 104 104
FT CARBOHD 177 177
FT CARBOHD 216 216
FT CARBOHD 232 232
FT CARBOHD 246 246
FT CARBOHD 271 271
SQ SEQUENCE 372 AA; 42188 MW; 6EA991802A2D3643 CRC64;

Query Match 95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1.8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEIINHTNCMDVGYGPOCQ 33
```

```

Db      160 CQPMSCSGHGEVCIINNTCNCVDGYGPQCQ 192
|||||
RESULT  4
LEML_PAPHA STANDARD: PRT: 372 AA.
AC      028768;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE      molecule-1) (LAM-1) (Leukocyte endothelial cell adhesion molecule 1)
DE      (LECAM1) (CD62L).
GN      SELL.
OS      Papio hamadryas (Hamadryas baboon).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC      Cercopithecinae; Papio.
OX      NCBI_TaxID=9557;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      MEDLINE=97128794; PubMed=8973334;
RA      Tsunushita N., Fu H., Berg E.L.;
RT      PCR cloning of the cDNA encoding baboon L-selectin.";
RL      Gene 181:219-220(1996).
CC      -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC      OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC      VENULES IN PERIPHERAL LYMPH NODES.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC      -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U52074; ABA0903.1; -.
DR      HSSP: P14151; 1KJB.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR002396; Selectin.
DR      InterPro: IPR000436; Sush1_SCR_CCP.
DR      InterPro: IPR001304; lectin_c.
DR      Pfam: PF00008; EGF_1.
DR      Pfam: PF00059; lectin_c; 1.
DR      PRINTS: PR00343; SELECTIN.
DR      SMART: SM00032; CCP; 2.
DR      SMART: SM00034; CLECT; 1.
DR      SMART: SM00181; EGF_1.
DR      PROSITE: PS00022; EGF_1; 1.
DR      PROSITE: PS01186; EGF_2; 1.
DR      PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR      PROSITE: PS50041; C-TYPE_LECTIN_2; 1.
DR      Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW      Selectin; Signal; Sush1; Repeat.
FT      SIGNAL 1 28
FT      PROPEP 29 38
FT      CHAIN 39 372
FT      DOMAIN 39 332
FT      TRANSMEM 333 355
FT      DOMAIN 356 372
FT      DOMAIN 55 155
FT      DOMAIN 156 182
FT      DOMAIN 196 255
FT      DOMAIN 258 317
FT      DISULFID 57 155

```

```

FT      DISULFID 128 147
FT      DISULFID 160 171
FT      DISULFID 165 180
FT      DISULFID 182 191
FT      DISULFID 197 241
FT      DISULFID 227 254
FT      DISULFID 259 303
FT      DISULFID 289 316
FT      CARBOHYD 60 60
FT      CARBOHYD 104 104
FT      CARBOHYD 177 177
FT      CARBOHYD 226 226
FT      CARBOHYD 232 232
FT      CARBOHYD 246 246
FT      CARBOHYD 271 271
SQ      SEQUENCE 372 AA; 42091 MW; 64E7BD5AC549D69 CRC64;

Query Match      95.3%; Score 205; DB 1; Length 372;
Best Local Similarity 97.0%; Pred. No. 1,8e-18;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 CQPMSCSGHGEVCIINNTCNCVDGYGPQCQ 33
Db      160 CQPMSCSGHGEVCIINNTCNCVDGYGPQCQ 192
|||||
RESULT  5
LEML_PONPY STANDARD: PRT: 372 AA.
AC      Q95235;
DT      01-NOV-1997 (Rel. 35, Created)
DT      01-NOV-1997 (Rel. 35, Last sequence update)
DT      01-NOV-1997 (Rel. 35, Last annotation update)
DE      L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE      molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE      (LECAM1) (CD62L).
GN      SELL.
OS      Pongo pygmaeus (Orangutan).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pongo.
OX      NCBI_TaxID=9600;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA      Tsunushita N.;
RT      Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.
RL      FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC      OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC      VENULES IN PERIPHERAL LYMPH NODES.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC      -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation-
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL: U73729; AAB18247.1; -.
DR      HSSP: P14151; 1KJB.
DR      InterPro: IPR000561; EGF-like.
DR      InterPro: IPR002396; Selectin.
DR      InterPro: IPR000436; Sush1_SCR_CCP.
DR      InterPro: IPR001304; lectin_c.
DR      Pfam: PF00008; EGF_1.
DR      Pfam: PF00059; lectin_c; 1.
DR      Pfam: PF00084; sush1; 2.

```

DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
DR PROSITE; PS0041; C-TYPE LECTIN\_2; 1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
Cell lectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28 BY SIMILARITY.  
FT PROPEP 29 38 BY SIMILARITY.  
FT CHAIN 39 372 L-SELECTIN.  
FT TRANSMEM 39 332 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 333 355 POTENTIAL.  
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).  
FT DOMAIN 156 192 EGF-LIKE.  
FT DOMAIN 196 255 SUSHI 1.  
FT DOMAIN 258 317 SUSHI 2.  
FT DISULFID 57 155 BY SIMILARITY.  
FT DISULFID 128 147 BY SIMILARITY.  
FT DISULFID 160 171 BY SIMILARITY.  
FT DISULFID 165 180 BY SIMILARITY.  
FT DISULFID 182 191 BY SIMILARITY.  
FT DISULFID 197 241 BY SIMILARITY.  
FT DISULFID 227 254 BY SIMILARITY.  
FT DISULFID 259 303 BY SIMILARITY.  
FT DISULFID 289 316 BY SIMILARITY.  
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 372 AA; 42118 MW; 6517DD2213PF15E CRC64;

Query Match 95.3%; Score 205; DB 1; Length 372;  
Best Local Similarity 97.0%; Pred. No. 1.8e-18;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
OY 1 COPWSCSGHCEVEIINHTCNCVDVGYGPOCQ 33  
DB 160 COPWSCSGHCEVEIINHTCNCVDVGYGPOCQ 192  
RESULT 6  
ID LEM1\_BOVIN STANDARD; PRT; 370 AA.  
AC P98131;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)  
DE (LECAM1) (CD62L).  
GN SELL.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92164727; PubMed=1371468;  
RA Walcheck B., White M., Kurk S., Kishimoto T.K., Juttla M.A.;  
RT "Characterization of the bovine peripheral lymph node homing  
receptor: a lectin cell adhesion molecule (LECAM).";  
RL Eur. J. Immunol. 22:469-476(1992).  
RN [2]

RP SEQUENCE FROM N.A.  
RX MEDLINE=94055053; PubMed=7694420;  
RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;  
RT "Bovine L-selectin: a peripheral lymphocyte homing receptor.";  
RL Vet. Immunol. Immunopathol. 37:201-215(1993).  
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE  
OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
VENUES IN PERIPHERAL LYMPH NODES.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL; X62882; CAA44676.1; .  
CC HSSP; P14151; 1KJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00084; sushi; 2.  
DR Pfam; PF00059; lectin\_c; 1.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF\_2; 1.  
DR PROSITE; PS00615; C-TYPE LECTIN\_1; 1.  
DR PROSITE; PS0041; C-TYPE LECTIN\_2; 1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
Cell lectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28 POTENTIAL.  
FT PROPEP 29 38 POTENTIAL.  
FT CHAIN 39 370 L-SELECTIN.  
FT TRANSMEM 39 333 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 334 354 POTENTIAL.  
FT DOMAIN 355 370 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).  
FT DOMAIN 156 192 EGF-LIKE.  
FT DOMAIN 196 255 SUSHI 1.  
FT DOMAIN 258 317 SUSHI 2.  
FT DISULFID 57 155 BY SIMILARITY.  
FT DISULFID 128 147 BY SIMILARITY.  
FT DISULFID 160 171 BY SIMILARITY.  
FT DISULFID 165 180 BY SIMILARITY.  
FT DISULFID 182 191 BY SIMILARITY.  
FT DISULFID 197 241 BY SIMILARITY.  
FT DISULFID 227 254 BY SIMILARITY.  
FT DISULFID 259 303 BY SIMILARITY.  
FT DISULFID 289 316 BY SIMILARITY.  
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 77 77 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 308 308 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 370 AA; 41971 MW; 92168B8116AE9228 CRC64;

Query Match

88.8%; Score 191; DB 1; Length 370;



DT 01-JUL-1993 (Rel. 26, last sequence update)  
 DT 01-FEB-1996 (Rel. 33, last annotation update)  
 DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (LY-22) (Lymphocyte surface MEL-14 antigen)  
 DE (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).  
 GN SEL OR LNH OR LY-22.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92329548; PubMed=1378303;  
 RA Matenabe T., Song Y., Hirayama Y., Tamatani T., Kuida K., Miyasaka M.;  
 RT "Sequence and expression of a rat cDNA for LECAM-1.";  
 RL Biochim. Biophys. Acta 1131:321-324(1992).  
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: D10831; BAA01613.1; -  
 DR PIR: S23936; S23936.  
 DR HSSP: P14151; 1KJ9.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 2.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 2.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE\_LLECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin; Selectin; Signal; Sush1; Repeat.  
 KW SIGNAL  
 FT CHAIN 1 28  
 FT PROPEP 29 38  
 FT CHAIN 39 372  
 FT DOMAIN 39 372  
 FT TRANSMEM 333 355  
 FT DOMAIN 333 372  
 FT DOMAIN 356 372  
 FT DOMAIN 155 155  
 FT DOMAIN 156 192  
 FT DOMAIN 196 255  
 FT DOMAIN 258 317  
 FT DISULFID 37 155  
 FT DISULFID 128 147  
 FT DISULFID 160 171  
 FT DISULFID 165 180  
 FT DISULFID 182 191  
 FT DISULFID 197 241  
 FT DISULFID 227 254  
 FT DISULFID 259 303  
 FT DISULFID 289 316  
 FT CARBOHYD 60  
 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 278 278 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 372 AA; 42441 MW; 3888AE0F1EAD191A CRC64;  
 SQ SEQUENCE  
 Query Match 74.0%; Score 159; DB 1; Length 372;  
 Best Local Similarity 78.8%; Pred. No. 9e-13;  
 Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 COPWCSGHCCEVEIINHHNCNDGVYGPQC 33  
 DB 160 COPESCRHCEVEITNNHTICDPGYGPQC 192  
 RESULT 9  
 ID LEM2\_CANFA STANDARD; PRT; 611 AA.  
 AC P33730;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, last sequence update)  
 DT 01-FEB-1996 (Rel. 33, last annotation update)  
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1) (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2) (CD62E).  
 GN SELE.  
 OS Canis familiaris (Dog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
 NCBI\_TaxID=9615;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=jugular vein;  
 RA Manning A.M., Lane C.L., Auchampach J.A., Kukielka G.L.,  
 RA Rosenbloom C.L., Anderson D.C.;  
 RL Submitted (XX-1993) to the EMBL/GenBank/DBI databases.  
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY ELAM-1 IS SIAYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF GLYCOPROTEINS).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L23087; AAA30843.1; -  
 DR HSSP: P16581; IESL.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 6.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 6.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 2.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.

DR PROSITE: PS50041; C-TYPE LECTIN 2; 1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 22  
FT CHAIN 23 611  
FT DOMAIN 23 557  
FT TRANSMEM 558 579  
FT DOMAIN 580 611  
FT DOMAIN 39 139  
FT DOMAIN 140 176  
FT DOMAIN 180 239  
FT DOMAIN 242 301  
FT DOMAIN 304 364  
FT DOMAIN 367 427  
FT DOMAIN 430 490  
FT DOMAIN 493 549  
FT DISULFID 41 139  
FT DISULFID 112 131  
FT DISULFID 144 155  
FT DISULFID 149 164  
FT DISULFID 166 175  
FT DISULFID 181 225  
FT DISULFID 211 238  
FT DISULFID 243 287  
FT DISULFID 273 300  
FT DISULFID 305 350  
FT DISULFID 336 363  
FT DISULFID 368 413  
FT DISULFID 399 426  
FT DISULFID 431 476  
FT DISULFID 462 489  
FT DISULFID 494 535  
FT DISULFID 521 548  
FT CARBOHYD 526 548  
FT CARBOHYD 161 161  
FT CARBOHYD 204 204  
FT CARBOHYD 266 266  
FT CARBOHYD 313 313  
FT CARBOHYD 333 333  
FT CARBOHYD 528 528  
SQ SEQUENCE 611 AA; 66314 MM; 35DA9E3DF225E4F6 CRC64;  
  
Query Match 60.9%; Score 131; DB 1; Length 611;  
Best Local Similarity 60.6%; Pred. No. 4.2e-09;  
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;  
  
QY 1 CQWSCSGHGEVLEINNHCTCNDVGYGPOCQ 33  
Db 144 CPTSCSGHGEVLEINNHCTCNDVGYGPOCQ 176  
  
RESULT 10  
LEM2\_HUMAN STANDARD; PRT; 610 AA.  
ID LEM2\_HUMAN  
AC P16581; P16111;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
DE (CD62e).  
GN Homo sapiens (Human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90175359; PubMed=1689848;  
RA Hession C., Osborn L., Goff D., Chi-Rosso G., Vassallo C.,  
RA Pasek M., Pittack C., Tizard R., Goelz S., McCarthy K., Hopple S.,  
RA Lobb R.;

RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning  
RT and functional interactions.";  
RT Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89162047; PubMed=2466335;  
RA Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.;  
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for  
RT neutrophils related to complement regulatory proteins and lectins.";  
RL Science 243:1160-1165(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91115870; PubMed=1703529;  
RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,  
RA Gimbrone M.A. Jr., Bevilacqua M.P.;  
RT "Structure and chromosomal location of the gene for endothelial-  
RT leukocyte adhesion molecule 1.";  
RL J. Biol. Chem. 266:2466-2473(1991).  
RN [4]  
RP LIGAND.  
RX MEDLINE=91068005; PubMed=1701274;  
RA Phillips M.L., Nudelman E., Gaeta F.C., Perez M., Singhal A.K.,  
RA Hakomori S., Paulson J.C.;  
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate  
RT ligand, sialyl-Lex.";  
RL Science 250:1130-1132(1990).  
RN [5]  
RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.  
RX MEDLINE=93202275; PubMed=7681016;  
RA Mills A.;  
RT "Modelling the carbohydrate recognition domain of human E-selectin.";  
RL FEBS Lett. 319:5-11(1993).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.  
RX MEDLINE=94150646; PubMed=7509040;  
RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,  
RA Huang K.-S., Presky D.H., Familletti P.C., Wolitzky B.A., Burns D.K.;  
RT "Insight into E-selectin/ligand interaction from the crystal  
RT structure and mutagenesis of the lec/EGF domains.";  
RL Nature 367:532-538(1994).  
RN [7]  
RP VARIANT ARG-149.  
RX MEDLINE=95179107; PubMed=7533025;  
RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,  
RA Schulte K.L., Glaser C., Rohde K., Baumann G., Speer A.;  
RT "E-selectin polymorphism and atherosclerosis: an association study.";  
RL Hum. Mol. Genet. 3:1935-1937(1994).  
RN [8]  
RP VARIANT ARG-149.  
RX MEDLINE=99134508; PubMed=9933738;  
RA Ye S.-Q., Usher D., Virgil D., Zhang L.-Q., Yochim S.E., Gupta R.;  
RT "A pecti polymorphism detects the mutation of serine-128 to arginine in  
RT CD 62e gene - a risk factor for coronary artery disease.";  
RL J. Biomed. Sci. 6:18-21(1999).  
RN [9]  
RP FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
RP MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
RP ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
RP POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
RP GLYCOLIPIDS).  
RN [10]  
RP SUBCELLULAR LOCATION: Type I membrane protein.  
RN [11]  
RP POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A  
RN HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY  
RN HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH  
RN ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN  
RN UNSELECTED POPULATION (SER-149).  
RN [12]  
RP SIMILARITY: TO OTHER SELECTINS/LICAMS.  
RN [13]  
RP SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
RN [14]  
RP SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.  
RN [15]  
RP DATABASE: NAME=PROW; NOTE=CD guide CD62e entry;  
RN WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62e.htm"  
RN [16]  
RN This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on how  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

[illegible]

```

FT CARBOHYD 199 199 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 332 332 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 503 503 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 527 527 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARIANT 130 130 C -> W (IN DBSNP:5360) .
FT VARIANT 149 149 /FTID=VAR.011790.
FT VARIANT 149 149 S -> R (ASSOCIATED WITH A RISK FACTOR FOR CAD) .
FT VARIANT 295 295 /FTID=VAR.004191.
FT VARIANT 295 295 E -> K (IN DBSNP:5364) .
FT VARIANT 421 421 /FTID=VAR.011791.
FT VARIANT 421 421 E -> Q (IN DBSNP:5366) .
FT VARIANT 468 468 /FTID=VAR.011792.
FT VARIANT 468 468 H -> Y (IN DBSNP:5368) .
FT VARIANT 575 575 /FTID=VAR.011793.
FT VARIANT 575 575 L -> F (IN DBSNP:5355) .
FT SEQUENCE 610 AA; 66655 MM; 7D43E3C0D1229229 CRC64;
SQ
Query Match 60.5%; Score 130; DB 1; Length 610;
Best Local Similarity 63.6%; Pred. No. 5.6e-09;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
OY 1 CQPMSCGSGHCEVEIITNNHNTCMCDVGYGPOCQ 33
DB 143 CMTSCSGHGEVEITNNYTKCKDGFSGLACE 175
RESULT 11
LEM2_RAT
ID LEM2_RAT STANDARD; PRT; 549 AA.
AC P98105.
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE E-selectin precursor (endothelial leukocyte adhesion molecule 1)
DE ELAM-1) (leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE (CD62e).
GN SELE OR ELAM-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;
RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY
CC ELAM-1 IS STALYNE-LWIS X (ALPHA11->)FUCOSYLATED DERIVATIVES OF
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF
CC GLYCOPOLIPIDS).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
EMBL; L25527; AAA41113.1; -.
OR HSSP; P16581; IKA.

```

DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c\_1.  
 DR Pfam: PF00084; sushi\_5.  
 DR SMART: SM00032; CCP\_5.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE\_LLECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sushi; Repeat.  
 FT SIGNAL 1 21 POTENTIAL.  
 FT CHAIN 1 21 E-SELECTIN.  
 FT DOMAIN 22 549 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 22 494 POTENTIAL.  
 FT DOMAIN 495 516 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 517 549 C-TYPE LECTIN (SHORT FORM).  
 FT DOMAIN 138 138 EGF-LIKE.  
 FT DOMAIN 139 175 EGF-LIKE.  
 FT DOMAIN 179 239 SUSHI 1.  
 FT DOMAIN 242 301 SUSHI 2.  
 FT DOMAIN 304 364 SUSHI 3.  
 FT DOMAIN 367 427 SUSHI 4.  
 FT DOMAIN 430 486 SUSHI 5.  
 FT DISULFID 40 138 BY SIMILARITY.  
 FT DISULFID 111 130 BY SIMILARITY.  
 FT DISULFID 143 154 BY SIMILARITY.  
 FT DISULFID 148 163 BY SIMILARITY.  
 FT DISULFID 165 174 BY SIMILARITY.  
 FT DISULFID 180 225 BY SIMILARITY.  
 FT DISULFID 210 238 BY SIMILARITY.  
 FT DISULFID 243 287 BY SIMILARITY.  
 FT DISULFID 273 300 BY SIMILARITY.  
 FT DISULFID 305 350 BY SIMILARITY.  
 FT DISULFID 336 363 BY SIMILARITY.  
 FT DISULFID 413 433 BY SIMILARITY.  
 FT DISULFID 399 426 BY SIMILARITY.  
 FT DISULFID 431 472 BY SIMILARITY.  
 FT DISULFID 458 485 BY SIMILARITY.  
 FT CARBOHYD 25 25 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 266 266 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 320 320 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 333 333 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 465 465 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 549 AA; 60079 MW; 85CEBCEB7B0144C8 CRC64;

Query Match 59.1%; Score 127; DB 1; Length 549;  
 Best Local Similarity 62.5%; Pred. No. 1.2e-08;  
 Matches 20; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

OY 1 CQPMSCGHEGCEVITNNHNCNDVGYGPOC 32  
 DB 143 CTNTSCGHEGCEVITNNHNCNDVGYGPOC 174

RESULT 12  
 LEM3\_BOVIN  
 ID LEM3\_BOVIN STANDARD; PRT; 646 AA.  
 AC P42201;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)

DE (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).  
 GN SEMP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Capillary endothelium;  
 RX MEDLINE-93249394; PubMed=7683458;  
 RA Strubel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;  
 RT "Isolation and characterization of a bovine cDNA encoding a  
 functional homolog of human P-selectin.";  
 RL Blochm. Biophys. Res. Commun. 192:338-344(1993).  
 CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS  
 TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE  
 INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH  
 LEUCOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS  
 AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL  
 ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO  
 THE CELL SURFACE.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS; BOVINE P-LECTIN LACKS  
 THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L12041; AAA30743.1; -.  
 DR HSSP: P16109; 1FSB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c\_1.  
 DR Pfam: PF00084; sushi\_5.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 6.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE\_LLECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sushi; Repeat.  
 FT SIGNAL 1 41 BY SIMILARITY.  
 FT CHAIN 1 41 P-SELECTIN.  
 FT DOMAIN 42 646 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 42 587 POTENTIAL.  
 FT DOMAIN 588 611 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 612 646 C-TYPE LECTIN (SHORT FORM).  
 FT DOMAIN 58 158 EGF-LIKE.  
 FT DOMAIN 159 195 SUSHI 1.  
 FT DOMAIN 199 258 SUSHI 2.  
 FT DOMAIN 261 320 SUSHI 3.  
 FT DOMAIN 323 382 SUSHI 4.  
 FT DOMAIN 385 444 SUSHI 5.  
 FT DOMAIN 457 516 SUSHI 6.  
 FT DOMAIN 519 578 SUSHI 7.  
 FT DISULFID 60 158 BY SIMILARITY.  
 FT DISULFID 131 150 BY SIMILARITY.  
 FT DISULFID 163 174 BY SIMILARITY.



```
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 458 502 BY SIMILARITY.
FT DISULFID 488 515 BY SIMILARITY.
FT DISULFID 520 564 BY SIMILARITY.
FT DISULFID 550 577 BY SIMILARITY.
FT CARBOHYD 48 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 532 532 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 539 539 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 557 557 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 634 637 ENDOCYTOSIS SIGNAL (PROBABLE).
SQ SEQUENCE 646 AA: 71229 MW: 573912A4627A6ACA CRC64:
```

```
Query Match 59.1%; Score 127; DB 1; Length 646;
Best Local Similarity 57.6%; Pred. No. 1.4e-08;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHCEVEIINHTCNDVGYRQPCQ 33
DB 163 CQPMSCSGHCEVEIINHTCNDVGYRQPCQ 195
```

```
RESULT 13
LEM3_SHEEP
ID LEM3_SHEEP STANDARD: PRT: 769 AA.
AC P98109;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE (CD62p) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN SELP.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Heart;
RA Burns S.A., Neufeld E.J., Donady J.J.;
RL Submitted (JUL-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entitles requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L34270; AAB59261.1; -.
DR HSSP: P16109; IFSB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_c.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; Lectin_c; 1.
DR Pfam: PF00084; sushi; 8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 32
FT CHAIN 33 769
FT DOMAIN 33 717
FT TRANSMEM 718 734
FT DOMAIN 735 759
FT DOMAIN 159 195
FT DOMAIN 199 258
FT DOMAIN 261 320
FT DOMAIN 323 382
FT DOMAIN 385 444
FT DOMAIN 447 506
FT DOMAIN 509 568
FT DOMAIN 580 639
FT DOMAIN 642 701
FT DISULFID 60 158
FT DISULFID 131 150
FT DISULFID 163 174
FT DISULFID 168 183
FT DISULFID 185 194
FT DISULFID 200 244
FT DISULFID 230 257
FT DISULFID 262 306
FT DISULFID 292 319
FT DISULFID 324 368
FT DISULFID 381 430
FT DISULFID 386 443
FT DISULFID 416 462
FT DISULFID 448 492
FT DISULFID 478 505
FT DISULFID 510 554
FT DISULFID 540 567
FT DISULFID 581 625
FT DISULFID 611 638
FT DISULFID 643 687
FT DISULFID 673 700
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT SITE 757 760
FT VARIANT 556 566 S->T.
FT VARIANT 579 579 L->V.
SQ SEQUENCE 769 AA: 84317 MW: 23842575D6FAB15 CRC64:
```









DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
 DE (ELAM-1) (leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
 DE (CD62E).  
 GN SELE.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 ON NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Adrenal gland;  
 RX MEDLINE=93382537; PubMed=7690465;  
 RA Nguyen M., Strubel N.A., Bischoff J.;  
 RT "A role for sialyl Lewis-X/A glycoconjugates in capillary  
 RT morphogenesis.";  
 RT Nature 365:267-269(1993).  
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
 CC GLYCOLIPIDS).  
 CC -1- FUNCTION: INVOLVED IN CAPILLARY MORPHOGENESIS BY BINDING TO A  
 CC BOVINE CAPILLARY ENDOTHELIAL (BCE) CELL SIALYL LEWIS X AND/OR  
 CC SIALYL LEWIS A-CONTAINING LIGAND (PROBABLY).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS. BOVINE E-LECTIN LACKS  
 CC THE HUMAN SUSHI-4 AND 5 EQUIVALENTS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L12039; AAA02991.1; -  
 DR HSSP: P16581; IESL.  
 DR Interpro: IPR0000561; EGF-like.  
 DR Interpro: IPR000742; EGF-2.  
 DR Interpro: IPR002396; Selectin.  
 DR Interpro: IPR000436; Sush1\_SCR\_CCP.  
 DR Interpro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 4.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 4.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE LECTIN\_2; 1.  
 DR PROSITE: PS50041; C-TYPE LECTIN\_1; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sush1; Repeat.  
 FT CHAIN 1 22  
 FT SIGNAL 23 485  
 FT DOMAIN 23 430  
 FT TRANSMEM 431 453  
 FT DOMAIN 454 485  
 FT DOMAIN 39 139  
 FT DOMAIN 140 176  
 FT DOMAIN 180 238  
 FT DOMAIN 241 300  
 FT DOMAIN 303 363  
 FT DOMAIN 366 422  
 FT DISULFID 41 139  
 FT DISULFID 112 131

FT DISULFID 144 155 BY SIMILARITY.  
 FT DISULFID 149 164 BY SIMILARITY.  
 FT DISULFID 166 175 BY SIMILARITY.  
 FT DISULFID 181 224 BY SIMILARITY.  
 FT DISULFID 210 237 BY SIMILARITY.  
 FT DISULFID 242 286 BY SIMILARITY.  
 FT DISULFID 272 299 BY SIMILARITY.  
 FT DISULFID 304 349 BY SIMILARITY.  
 FT DISULFID 335 362 BY SIMILARITY.  
 FT DISULFID 367 408 BY SIMILARITY.  
 FT DISULFID 394 421 BY SIMILARITY.  
 FT CARBOHYD 61 61 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 265 265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 312 312 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 316 316 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 379 379 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 401 401 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 485 AA; 53200 MW; AE931C9B521E3904 CRC64; (POTENTIAL).  
 SO  
 Query Match 54.9%; Score 118; DB 1; Length 485;  
 Best Local Similarity 54.5%; Pred. No. 1,4e-07;  
 Matches 18; Conservative 4; Mismatches 11; Indels 0; Gaps 0;  
 Oy 1 CQPMSCGSGHCEVCIINNHHCNCDEVGYGPQCQ 33  
 Db 144 CNPFCGSHGECVETINNYTCQCHPFGKJKE 176  
 ID LEM2\_RABIT STANDARD; PRT; 551 AA.  
 AC P27113;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
 DE (ELAM-1) (leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
 DE (CD62E).  
 GN SELE.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 ON NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=92189729; PubMed=1372169;  
 RA Larian J.D., Tsang T.C., Rumberger J.M., Burns D.K.;  
 RT "Characterization of cDNA and genomic sequences encoding rabbit  
 RT ELAM-1: conservation of structure and functional interactions with  
 RT leukocytes.";  
 RL DNA Cell Biol. 11:149-162(1992).  
 CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
 CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
 CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
 CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
 CC GLYCOLIPIDS).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- INDUCTION: BY CYTOKINES.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 258 SUSH1 1.
FT DOMAIN 261 320 SUSH1 2.
FT DOMAIN 323 382 SUSH1 3.
FT DOMAIN 385 444 SUSH1 4.
FT DOMAIN 447 506 SUSH1 5.
FT DOMAIN 509 568 SUSH1 6.
FT DOMAIN 579 638 SUSH1 7.
FT DOMAIN 641 700 SUSH1 8.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 448 492 BY SIMILARITY.
FT DISULFID 478 505 BY SIMILARITY.
FT DISULFID 510 554 BY SIMILARITY.
FT DISULFID 540 567 BY SIMILARITY.
FT DISULFID 580 624 BY SIMILARITY.
FT DISULFID 610 637 BY SIMILARITY.
FT DISULFID 642 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT CARBOHYD 54 45 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 107 107 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 456 456 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 745 745 PALMITATE (BY SIMILARITY).
FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).
SQ SEQUENCE 768 AA: 83517 MW: 26FD7E8A5F3F1316 CRC64;

```

Query Match 50.7%; Score 109; DB 1; Length 768;  
 Best Local Similarity 48.5%; Pred. No. 2,8e-06;  
 Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

```

QY 1 COPWSCSGHGEVEIINHTCNDVGYYGQCQ 33
    11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
Db 163 CQMSKNSGCRIFETISYTCSPGYGPECE 195

```

```

RESULT 21
ID FA12_CAVPO STANDARD; PRT: 603 AA.
AC 004962;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (HAF) (Fragment).
GN F12.
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 19-37; 318-332 AND 359-373.
RC TISSUE=Liver;
RA MEDLINE=93003367; PubMed=1390917;
RA Sema U., Yamamoto T., Kunisada T., Shibuya Y., Tanase S.,
RA Kambara T., Okabe H.;

```

```

RT "Primary structure of guinea-pig Hageman factor: sequence around the
RT cleavage site differs from the human molecule.";
RL Blochim. Biophys. Acta 1159:113-121(1992).
CC -1- FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.
CC -1- CATALYTIC ACTIVITY: Cleaves selectively Arg-1-ile bonds in factor
CC VII to form factor VIIa and factor XI to form factor XIa.
CC -1- MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMK KININOGEN FORM A
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST
CC TO ALPHA-FACTOR XIIA AND THEN TO BETA-FACTOR XIIA. ALPHA-FACTOR
CC XIIA ACTIVATES FACTOR XI TO FACTOR XIa.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE
CC TRYPSIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68615; CAA48600.1; -.
DR HSSP: P00750; IRTF.
DR MEROPS: S01.211; -.
DR InterPro: IPR000561; EGF-1-like.
DR InterPro: IPR000742; EGF 2.
DR InterPro: IPR001881; EGF Ca.
DR InterPro: IPR000562; FN_Type_II.
DR InterPro: IPR000001; Kringle.
DR InterPro: IPR001254; Trypsin.
DR InterPro: IPR000083; fibronectin_type_1.
DR Pfam: PF00008; EGF_2.
DR Pfam: PF00039; fn1. 1.
DR Pfam: PF00040; fn2. 1.
DR Pfam: PF00051; kringle. 1.
DR Pfam: PF00089; trypsin. 1.
DR ProDom: PD000995; FN_Type_II; 1.
DR SMART: SM00181; EGF; 2.
DR SMART: SM00058; FN1; 1.
DR SMART: SM00059; FN2; 1.
DR SMART: SM00130; KR; 1.
DR SMART: SM00020; tryp_Spec; 1.
DR PROSITE: PS00022; EGF_1; 2.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS01253; FIBRONECTIN_1; 1.
DR PROSITE: PS00023; FIBRONECTIN_2; 1.
DR PROSITE: PS00021; KRINGLE_1; 1.
DR PROSITE: PS00070; KRINGLE_2; 1.
DR PROSITE: PS50240; TRYPsin_DOM; 1.
DR PROSITE: PS00134; TRYPsin_HIS; 1.
DR PROSITE: PS00135; TRYPsin_SER; 1.
KW Glycoprotein; Blood coagulation; Plasma; Kringle; Serine protease;
KW Hydrolyase; Fibrinolysis; EGF-like domain; Repeat; Zymogen; Signal.
FT NON_TER 1 1
FT SIGNAL 1 18
FT CHAIN 19 358 ALPHA-FACTOR XIIA HEAVY CHAIN.
FT CHAIN 359 603 ALPHA-FACTOR XIIA LIGHT CHAIN.
FT DOMAIN 46 87 FIBRONECTIN TYPE-II.
FT DOMAIN 93 130 EGF-LIKE 1.
FT DOMAIN 132 172 FIBRONECTIN TYPE-I.
FT DOMAIN 173 209 EGF-LIKE 2.
FT DOMAIN 216 294 KRINGLE.
FT DOMAIN 312 342 PRO-RICH.
FT DOMAIN 359 603 SERINE PROTEASE.
FT ACT_SITE 398 398 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 447 447 CHANGE RELAY SYSTEM (BY SIMILARITY).

```



```
FT ACT_SITE 551 551 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT DISULFID 97 109 BY SIMILARITY.
FT DISULFID 103 118 BY SIMILARITY.
FT DISULFID 120 129 BY SIMILARITY.
FT DISULFID 134 162 BY SIMILARITY.
FT DISULFID 160 169 BY SIMILARITY.
FT DISULFID 177 188 BY SIMILARITY.
FT DISULFID 182 197 BY SIMILARITY.
FT DISULFID 199 208 BY SIMILARITY.
FT DISULFID 216 294 BY SIMILARITY.
FT DISULFID 237 276 BY SIMILARITY.
FT DISULFID 265 289 BY SIMILARITY.
FT DISULFID 345 472 BY SIMILARITY.
FT DISULFID 383 399 BY SIMILARITY.
FT DISULFID 391 461 BY SIMILARITY.
FT DISULFID 422 425 BY SIMILARITY.
FT DISULFID 488 557 BY SIMILARITY.
FT DISULFID 520 536 BY SIMILARITY.
FT DISULFID 547 578 BY SIMILARITY.
FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 270 270 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 419 419 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 603 AA; 66795 MW; 48DC6B946FB9BD59 CRC64;

Query Match 44.2%; Score 95; DB 1; Length 603;
Best Local Similarity 44.1%; Pred. No. 0.00012;
Matches 15; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEIINHTNCNDVGYGPQCOL 34
DB 177 CQTNPCNLNGRCLEVEGHHLCDCPKMYTGPFDL 210

RESULT 22
DILI_HUMAN STANDARD; PRT; 723 AA.
AC 000548:
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Delta-like protein 1 precursor (Drosophila Delta homolog 1) (Delta1)
DE (H-Delta-1).
CN DILI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBITaxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99180765; PubMed=10079256;
RA Gray G.E., Mann R.S., Mitsiadis E., Henrique D., Carcangiu M.L.,
RA Banks A., Leiman J., Ward D., Ish-Horowitz D., Artavanis-Tsakonas S.;
RT "Human ligands of the Notch receptor.";
RL Am. J. Pathol. 154:785-794(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Oda T., Chandrasekharappa S.C.;
RT "Human Delta 1 gene sequence.";
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN
CC MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS
CC UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
```

```
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL; AF003522; AAB61286.1; -.
DR EMBL; AF222310; AAG09716.1; -.
DR HSSP; P00740; 1EDM.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR001774; DSL.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF_II.
DR Pfam; PR01414; DSL; 1.
DR Pfam; PR00008; EGF; 8.
DR PRINTS; PR00010; EGFBL00D.
DR SMART; SM00051; DSL; 1.
DR SMART; SM00179; EGF_CA; 4.
DR SMART; SM00001; EGF_like; 4.
DR PROSITE; PS00010; ASX_HYDROXYL; 3.
DR PROSITE; PS01186; EGF_2; 8.
DR PROSITE; PS01187; EGF_CA; 1.
DR Signal; EGF-like domain; Repeat; Transmembrane; Developmental protein;
KW Differentiation; Glycoprotein.
FT SIGNAL 1 17
FT CHAIN 18 723
FT DOMAIN 18 545 DELTA-LIKE PROTEIN 1.
FT TRANSMEM 546 568 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 569 723 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 226 254 EGF-LIKE 1.
FT DOMAIN 257 285 EGF-LIKE 2.
FT DOMAIN 292 325 EGF-LIKE 3.
FT DOMAIN 332 363 EGF-LIKE 4.
FT DOMAIN 370 402 EGF-LIKE 5.
FT DOMAIN 409 440 EGF-LIKE 6.
FT DOMAIN 447 478 EGF-LIKE 7.
FT DOMAIN 485 516 EGF-LIKE 8.
FT DISULFID 226 237 BY SIMILARITY.
FT DISULFID 230 243 BY SIMILARITY.
FT DISULFID 245 254 BY SIMILARITY.
FT DISULFID 257 268 BY SIMILARITY.
FT DISULFID 263 274 BY SIMILARITY.
FT DISULFID 276 285 BY SIMILARITY.
FT DISULFID 292 304 BY SIMILARITY.
FT DISULFID 298 314 BY SIMILARITY.
FT DISULFID 316 325 BY SIMILARITY.
FT DISULFID 332 343 BY SIMILARITY.
FT DISULFID 337 352 BY SIMILARITY.
FT DISULFID 354 363 BY SIMILARITY.
FT DISULFID 370 381 BY SIMILARITY.
FT DISULFID 375 391 BY SIMILARITY.
FT DISULFID 393 402 BY SIMILARITY.
FT DISULFID 409 420 BY SIMILARITY.
FT DISULFID 414 429 BY SIMILARITY.
FT DISULFID 431 440 BY SIMILARITY.
FT DISULFID 447 467 BY SIMILARITY.
FT DISULFID 469 478 BY SIMILARITY.
FT DISULFID 485 496 BY SIMILARITY.
FT DISULFID 490 505 BY SIMILARITY.
FT DISULFID 507 516 BY SIMILARITY.
FT CARBOHYD 477 477 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 723 AA; 77956 MW; B4EC455FFA32A12B CRC64;

Query Match 43.3%; Score 93; DB 1; Length 723;
Best Local Similarity 38.2%; Pred. No. 0.00026;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCGHGECVEIINHTNCNDVGYGPQCOL 34
DB 332 CDPWSCGHGECVDLENSYSCICPPGFKICEL 365
```

RESULT 23  
NOTC\_XENLA STANDARD: PRT: 2524 AA.  
ID NOTC\_XENLA  
AC P21783;  
DT 01-MAY-1991 (Rel. 18, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neurogenic locus notch protein homolog precursor (XOTCH protein).  
GN XOTCH.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_Taxid=8335;  
RN [1]  
RX SEQUENCE FROM N.A.  
RX MEDLINE=90385285; PubMed=2402639;  
RA Coffman C., Harris W., Kintner C.;  
RT "Xotch, the Xenopus homolog of Drosophila notch.";  
RL Science, 249:1438-1441(1990).  
RM [2]  
RP REVISIONS TO 1759-1782.  
RA Kintner C.;  
RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.  
CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1 DEVELOPMENTAL STAGE: EXPRESSED ALMOST UNIFORMLY IN EARLY EMBRYOS.  
CC -1 SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.  
CC -1 SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.  
CC -1 SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.  
CC -1 SIMILARITY: CONTAINS 6 ANK REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
DR EMBL; M33874; AAB02039.1; -.  
DR PIR; A35844; A35844.  
DR HSSP; P00740; 1EDM.  
DR InterPro: IPR002110; ANK.  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000742; EGF\_2.  
DR InterPro: IPR001881; EGF\_CA.  
DR InterPro: IPR001438; EGF\_II.  
DR InterPro: IPR000800; Notch.  
DR Pfam; PF00023; ank; 6.  
DR Pfam; PF00008; EGF; 36.  
DR Pfam; PF00066; notch; 3.  
DR PRINTS; PR00010; EGFBL00D.  
DR PRINTS; PR01452; NOTCH.  
DR SMART; SM00248; ANK; 5.  
DR SMART; SM00179; EGF\_CA; 23.  
DR SMART; SM00001; EGF-like; 11.  
DR SMART; SM00004; NL; 2.  
DR PROSITE; PS50088; ANK\_REPEAT; 4.  
DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.  
DR PROSITE; PS00010; ASX\_HYDROXYL; 23.  
DR PROSITE; PS00022; EGF\_1; 34.  
DR PROSITE; PS01186; EGF\_2; 29.  
DR PROSITE; PS01187; EGF\_CA; 21.  
KW Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;  
KW Transmembrane; Signal; Glycoprotein.  
FT SIGNAL 1 19 POTENTIAL.  
FT CHAIN 20 2524 NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG.  
FT DOMAIN 20 1728 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 1729 1730 POTENTIAL.  
FT DOMAIN 1751 2524 CYTOPLASMIC (POTENTIAL).  
FT 20 57 EGF-LIKE 1.  
FT DOMAIN 58 99 EGF-LIKE 2.

FT DOMAIN 102 140 EGF-LIKE 3.  
FT DOMAIN 141 177 EGF-LIKE 4.  
FT DOMAIN 179 215 EGF-LIKE 5.  
FT DOMAIN 217 254 EGF-LIKE 6.  
FT DOMAIN 256 292 EGF-LIKE 7.  
FT DOMAIN 294 332 EGF-LIKE 8.  
FT DOMAIN 334 370 EGF-LIKE 9.  
FT DOMAIN 371 409 EGF-LIKE 10.  
FT DOMAIN 411 449 EGF-LIKE 11.  
FT DOMAIN 451 487 EGF-LIKE 12.  
FT DOMAIN 489 525 EGF-LIKE 13.  
FT DOMAIN 527 563 EGF-LIKE 14.  
FT DOMAIN 565 600 EGF-LIKE 15.  
FT DOMAIN 602 638 EGF-LIKE 16.  
FT DOMAIN 640 675 EGF-LIKE 17.  
FT DOMAIN 677 713 EGF-LIKE 18.  
FT DOMAIN 715 750 EGF-LIKE 19.  
FT DOMAIN 752 788 EGF-LIKE 20.  
FT DOMAIN 790 826 EGF-LIKE 21.  
FT DOMAIN 828 866 EGF-LIKE 22.  
FT DOMAIN 868 904 EGF-LIKE 23.  
FT DOMAIN 906 942 EGF-LIKE 24.  
FT DOMAIN 944 980 EGF-LIKE 25.  
FT DOMAIN 982 1018 EGF-LIKE 26.  
FT DOMAIN 1020 1056 EGF-LIKE 27.  
FT DOMAIN 1058 1094 EGF-LIKE 28.  
FT DOMAIN 1142 1180 EGF-LIKE 29.  
FT DOMAIN 1144 1180 EGF-LIKE 30.  
FT DOMAIN 1182 1218 EGF-LIKE 31.  
FT DOMAIN 1220 1264 EGF-LIKE 32.  
FT DOMAIN 1266 1304 EGF-LIKE 33.  
FT DOMAIN 1306 1346 EGF-LIKE 34.  
FT DOMAIN 1347 1383 EGF-LIKE 35.  
FT DOMAIN 1386 1424 EGF-LIKE 36.  
FT REPEAT 1441 1478 LIN/NOTCH 1.  
FT REPEAT 1479 1520 LIN/NOTCH 2.  
FT REPEAT 1521 1560 LIN/NOTCH 3.  
FT REPEAT 1561 1599 ANK 1.  
FT REPEAT 1624 1662 ANK 2.  
FT REPEAT 1663 1701 ANK 3.  
FT REPEAT 1702 1740 ANK 4.  
FT REPEAT 1741 1779 ANK 5.  
FT REPEAT 1780 1818 ANK 6.  
FT REPEAT 1819 1857 BY SIMILARITY.  
FT DISULFID 22 35 BY SIMILARITY.  
FT DISULFID 29 45 BY SIMILARITY.  
FT DISULFID 47 56 BY SIMILARITY.  
FT DISULFID 62 74 BY SIMILARITY.  
FT DISULFID 68 87 BY SIMILARITY.  
FT DISULFID 89 98 BY SIMILARITY.  
FT DISULFID 106 117 BY SIMILARITY.  
FT DISULFID 111 128 BY SIMILARITY.  
FT DISULFID 130 139 BY SIMILARITY.  
FT DISULFID 145 156 BY SIMILARITY.  
FT DISULFID 150 165 BY SIMILARITY.  
FT DISULFID 167 176 BY SIMILARITY.  
FT DISULFID 183 194 BY SIMILARITY.  
FT DISULFID 188 203 BY SIMILARITY.  
FT DISULFID 205 214 BY SIMILARITY.  
FT DISULFID 221 232 BY SIMILARITY.  
FT DISULFID 226 242 BY SIMILARITY.  
FT DISULFID 244 253 BY SIMILARITY.  
FT DISULFID 260 271 BY SIMILARITY.  
FT DISULFID 265 280 BY SIMILARITY.  
FT DISULFID 282 291 BY SIMILARITY.  
FT DISULFID 298 311 BY SIMILARITY.  
FT DISULFID 305 320 BY SIMILARITY.  
FT DISULFID 322 331 BY SIMILARITY.  
FT DISULFID 338 349 BY SIMILARITY.  
FT DISULFID 343 358 BY SIMILARITY.  
FT DISULFID 360 369 BY SIMILARITY.  
FT DISULFID 375 386 BY SIMILARITY.  
FT DISULFID 380 397 BY SIMILARITY.  
FT DISULFID 399 408 BY SIMILARITY.

```

FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 504 BY SIMILARITY.
FT DISULFID 498 513 BY SIMILARITY.
FT DISULFID 515 524 BY SIMILARITY.
FT DISULFID 531 542 BY SIMILARITY.
FT DISULFID 536 551 BY SIMILARITY.
FT DISULFID 553 562 BY SIMILARITY.
FT DISULFID 569 579 BY SIMILARITY.
FT DISULFID 574 588 BY SIMILARITY.
FT DISULFID 590 599 BY SIMILARITY.
FT DISULFID 606 617 BY SIMILARITY.
FT DISULFID 611 626 BY SIMILARITY.
FT DISULFID 628 637 BY SIMILARITY.
FT DISULFID 644 654 BY SIMILARITY.
FT DISULFID 649 663 BY SIMILARITY.
FT DISULFID 665 674 BY SIMILARITY.
FT DISULFID 681 692 BY SIMILARITY.
FT DISULFID 686 701 BY SIMILARITY.
FT DISULFID 703 712 BY SIMILARITY.
FT DISULFID 719 729 BY SIMILARITY.
FT DISULFID 724 738 BY SIMILARITY.
FT DISULFID 740 749 BY SIMILARITY.
FT DISULFID 756 767 BY SIMILARITY.
FT DISULFID 761 776 BY SIMILARITY.
FT DISULFID 778 787 BY SIMILARITY.
FT DISULFID 794 805 BY SIMILARITY.
FT DISULFID 799 814 BY SIMILARITY.
FT DISULFID 816 825 BY SIMILARITY.
FT DISULFID 832 843 BY SIMILARITY.
FT DISULFID 837 854 BY SIMILARITY.
FT DISULFID 856 865 BY SIMILARITY.
FT DISULFID 872 883 BY SIMILARITY.
FT DISULFID 877 892 BY SIMILARITY.
FT DISULFID 894 903 BY SIMILARITY.
FT DISULFID 910 921 BY SIMILARITY.
FT DISULFID 915 930 BY SIMILARITY.
FT DISULFID 932 941 BY SIMILARITY.
FT DISULFID 986 997 BY SIMILARITY.
FT DISULFID 991 1006 BY SIMILARITY.
FT DISULFID 1008 1017 BY SIMILARITY.
FT DISULFID 1024 1035 BY SIMILARITY.
FT DISULFID 1029 1044 BY SIMILARITY.
FT DISULFID 1046 1055 BY SIMILARITY.
FT DISULFID 1062 1073 BY SIMILARITY.
FT DISULFID 1067 1082 BY SIMILARITY.
FT DISULFID 1084 1093 BY SIMILARITY.
FT DISULFID 1100 1121 BY SIMILARITY.
FT DISULFID 1112 1130 BY SIMILARITY.
FT DISULFID 1132 1141 BY SIMILARITY.
FT DISULFID 1148 1159 BY SIMILARITY.
FT DISULFID 1153 1168 BY SIMILARITY.

```

Query Match 43.3%; Score 93; DB 1; Length 2524;  
 Best Local Similarity 45.5%; Pred. No. 0.00081;  
 Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSHGCEVEITNHTNCNDGYVGPCCO 33  
 DB 910 COPNPGHNGSCSDGIMFPCNCPAGRGPKCE 942

RESULT 24  
 YNXX3\_CABEL

ID YNXX3\_CABEL STANDARD; PRT: 3051 AA.  
 AC P34576;  
 DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1994 (Rel. 28, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

DE Hypothetical protein T20G5.3 in chromosome III (Fragment).
GN T20G5.3.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=BRISTOL N2;
RA Berns M., Smith A.;
RL Submitted (Mar-1994) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 45 EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 2 SEA DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 VFMA DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: Z30423; CAA83007.1; -.
CC PIR: S42373; S42373.
CC HSSP: P02468; 1TLE.
DR WormPep: T20G5.3; CE00478.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR000082; SEA.
DR InterPro: IPR002035; VFMA.
DR Pfam: PF00008; EGF; 30.
DR Pfam: PF01390; SEA; 2.
DR Pfam: PF00092; wva; 1.
DR SMART: SM00179; EGF_Ca; 1.
DR SMART: SM00001; EGF_Like; 38.
DR SMART: SM00320; SEA; 2.
DR SMART: SM00327; WVA; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; 28.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 5.
DR PROSITE: PS01187; EGF_Ca; 1.
DR PROSITE: PS50024; SEA; 2.
DR PROSITE: PS50334; WVA; 1.
KW Hypothetical protein; EGF-like domain; Repeat; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 2701
FT TRANSMEM 2702 2722
FT 2723 3051
FT 44
FT DOMAIN <1 44
FT 46 94
FT 100 144
FT 145 192
FT 194 245
FT 247 296
FT 313 354
FT 355 402
FT 404 452
FT 454 503
FT 514 590
FT 591 750
FT 751 794
FT 795 846
FT 847 892
FT 893 940
FT 941 990
FT 991 1039
FT 1040 1091
FT 1092 1144
FT 1145 1195
FT 1196 1245
FT 1246 1295
FT 1296

```



OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA MEDLINE=92316897; PubMed=1618739;  
RT Tada N., Sato M., Tsujimura A., Iwase R., Hashimoto-Gotoh T.;  
RL "Isolation and characterization of a mouse protein C cDNA.";  
RN J. Biochem. 111:491-495(1992).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129/SVJ;  
RA MEDLINE=98152576; PubMed=9493582;  
RT Jalbert L.R., Rosen E.D., Lissens A., Carmeliet P., Collen D.,  
RA Castellino F.J.;  
RL "Nucleotide structure and characterization of the murine gene encoding  
anticoagulant protein C.";  
RN Thromb. Haemost. 79:310-316(1998).  
[3]  
RP SEQUENCE OF 274-434 FROM N.A.  
RC STRAIN=BALB/C;  
RA MEDLINE=94318474; PubMed=8043441;  
RT Murakawa M., Okamura T., Kamura T., Kuroiwa M., Harada M., Niho Y.;  
RL "A comparative study of partial primary structures of the catalytic  
region of mammalian protein C.";  
RN Br. J. Haematol. 86:590-600(1994).  
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
and VIIa.  
CC -1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
TETRAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS  
REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
CC STRONGLY PROMOTED BY THROMBOMODULIN.  
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
CC -1- PTM: THE VITAMIN K-DEPENDENT ENZYMIC CARBOXYLATION OF SOME  
GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
SITE IS NECESSARY FOR THE RECOGNITION OF THE  
THROMBIN-THROMBOMODULIN COMPLEX.  
CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
TRIPESIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D10445; BAA01235.1; -;  
DR EMBL: AF034569; AAC33795.1; -;  
DR EMBL: D43755; BAA07812.1; -;  
DR PIR: JX0210; JX0210.  
DR HSSP: P04070; 1PCU.  
DR MEGS: S01.218; -;  
DR MGD: MGI:97771; PROC.  
DR InterPro: IPR000152; ASX\_hydroxyl.  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR002383; GLA\_blood.  
DR InterPro: IPR001254; Trypsin.  
DR InterPro: IPR000294; Vitk\_dep\_GLA.  
DR Pfam: PF00008; EGF\_2.  
DR Pfam: PF00594; gla; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.

DR PRINTS: PR00001; GLABLOOD.  
DR SMART: SM00179; EGF\_CA; 1.  
DR SMART: SM00001; EGF\_Like; 1.  
DR SMART: SM00069; GLA; 1.  
DR SMART: SM00020; TRYP\_Spc; 1.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01187; EGF\_CA; 1.  
DR PROSITE: PS00011; GLU-CARBOXYLATION; 1.  
DR PROSITE: PS50240; TRYPsin\_DOM; 1.  
DR PROSITE: PS00134; TRYPsin\_HIS; 1.  
DR PROSITE: PS00135; TRYPsin\_SER; 1.  
KW Blood coagulation; Glycoprotein; Serine protease;  
KW Gamma-carboxyglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
KW EGF-like domain; Repeat; Endothelial cell; Hydrolase; Signal.  
FT SIGNAL 1 33  
FT PROPEP 34 41  
FT CHAIN 42 196  
FT CHAIN 199 461  
FT PEPTIDE 199 212  
FT SITE 212 213  
FT DOMAIN 96 131  
FT DOMAIN 135 175  
FT DOMAIN 213 461  
FT MOD\_RES 47 47  
FT MOD\_RES 48 48  
FT MOD\_RES 55 55  
FT MOD\_RES 57 57  
FT MOD\_RES 60 60  
FT MOD\_RES 61 61  
FT MOD\_RES 66 66  
FT MOD\_RES 67 67  
FT MOD\_RES 70 70  
FT MOD\_RES 112 112  
FT ACT\_SITE 253 253  
FT ACT\_SITE 299 299  
FT ACT\_SITE 402 402  
FT DISULFID 58 63  
FT DISULFID 91 110  
FT DISULFID 100 105  
FT DISULFID 104 119  
FT DISULFID 121 130  
FT DISULFID 139 150  
FT DISULFID 146 159  
FT DISULFID 161 174  
FT DISULFID 162 319  
FT DISULFID 238 254  
FT DISULFID 373 387  
FT DISULFID 398 426  
FT CARBOHYD 214 290  
FT CARBOHYD 290 355  
FT CARBOHYD 355 355  
FT CONFLICT 328 328  
FT CONFLICT 393 393  
SQ SEQUENCE 461 AA; 51945 MW; 53FAAD085194D6E CRC64;

Query Match Score 87; DB 1; Length 461;  
Best Local Similarity 46.48; Pred. No. 0.00094;  
Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;  
QY 6 CSQHGEVETLNNHTCNCDVGYGYPQCQ 33  
| | | | | : : : | : | | | : | |

DB 104 CCGHCTCIGDGSFSCSKGMEKFCQ 131

RESULT 26

DL1L\_RAT

ID DL1L\_RAT STANDARD: PRT: 714 AA.

AC P97677;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Delta-like protein 1 precursor (Drosophila Delta homolog 1) (Delta).

GN DL1L.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.

OX NCBI\_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RL Disibio G., Hebshi L., Boulter J., Weimaster G.; Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: MAY BE INVOLVED IN CELL-TO-CELL COMMUNICATION IN MAMMALIAN EMBRYOS. MAY HAVE A ROLE IN CELLULAR INTERACTIONS UNDERLYING SOMITOGENESIS AND DEVELOPMENT OF THE NERVOUS SYSTEM (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.

CC -1- SIMILARITY: BELONGS TO THE DELTA/SERRATE/JAGGED FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation- the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: U78889; AAB37343.1; -

DR HSSP: P08709; I8P9; -

DR InterPro: IPR000152; Asx\_hydroxyl.

DR InterPro: IPR001774; DSL.

DR InterPro: IPR000561; EGF-Like.

DR InterPro: IPR000742; EGF\_2.

DR InterPro: IPR001881; EGF\_Ca.

DR InterPro: IPR001438; EGF\_IT.

DR Pfam: PF01414; DSL; 1.

DR Pfam: PF00008; EGF; 7.

DR PRINTS: PR00010; EGFBL00D.

DR SMART: SM000179; EGF\_CA; 4.

DR SMART: SM00001; EGF\_Like; 4.

DR PROSITE: PS00010; ASX\_HYDROXYL; 3.

DR PROSITE: PS00022; EGF\_1; 8.

DR PROSITE: PS01186; EGF\_2; 8.

DR PROSITE: PS01187; EGF\_CA; 2.

KW Signal: EGF-like domain; Repeat: Transmembrane; Developmental protein; Differentiation; Glycoprotein.

KW SIGNAL

FT 1 17

FT CHAIN

FT 18 714 DELTA-LIKE PROTEIN 1.

FT DOMAIN

FT 18 537 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM

FT 538 560 POTENTIAL.

FT DOMAIN

FT 561 714 CYTOPLASMIC (POTENTIAL).

FT DOMAIN

FT 225 253 EGF-LIKE 1.

FT DOMAIN

FT 256 284 EGF-LIKE 2.

FT DOMAIN

FT 291 324 EGF-LIKE 3.

FT DOMAIN

FT 331 362 EGF-LIKE 4.

FT DOMAIN

FT 369 401 EGF-LIKE 5.

FT DOMAIN

FT 408 439 EGF-LIKE 6.

FT DOMAIN

FT 446 477 EGF-LIKE 7.

FT DOMAIN

FT 484 515 EGF-LIKE 8.

FT DOMAIN

FT 515 536 BY SIMILARITY.

FT DISULFID

FT 225 236 BY SIMILARITY.

FT DISULFID

FT 229 242 BY SIMILARITY.

FT DISULFID

FT 244 253 BY SIMILARITY.

FT DISULFID

FT 256 267 BY SIMILARITY.

FT DISULFID 262 273 BY SIMILARITY.

FT DISULFID 275 284 BY SIMILARITY.

FT DISULFID 291 303 BY SIMILARITY.

FT DISULFID 297 313 BY SIMILARITY.

FT DISULFID 315 324 BY SIMILARITY.

FT DISULFID 331 342 BY SIMILARITY.

FT DISULFID 336 351 BY SIMILARITY.

FT DISULFID 353 362 BY SIMILARITY.

FT DISULFID 369 380 BY SIMILARITY.

FT DISULFID 374 390 BY SIMILARITY.

FT DISULFID 392 401 BY SIMILARITY.

FT DISULFID 408 419 BY SIMILARITY.

FT DISULFID 413 428 BY SIMILARITY.

FT DISULFID 430 439 BY SIMILARITY.

FT DISULFID 446 466 BY SIMILARITY.

FT DISULFID 468 477 BY SIMILARITY.

FT DISULFID 484 495 BY SIMILARITY.

FT DISULFID 489 504 BY SIMILARITY.

FT DISULFID 506 515 BY SIMILARITY.

FT CARBOHYD 476 476 N-LINKED (G1CNA... ) (POTENTIAL).

SO SEQUENCE 714 AA; 77378 MW; 4B8E2272BAEA27E CRC64;

Query Match 40.5%; Score 87; DB 1; Length 714;

Best Local Similarity 35.3%; Pred. No. 0.0014;

Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPSGSHGCVIITNNHCNDVGYRQCL 34

DB 331 CAPSPCRNGSGCTDLEDSYSCQCPGFGVCEL 364

RESULT 27

FBPL\_STRPU

ID FBPL\_STRPU STANDARD: PRT: 1064 AA.

AC P10079;

DT 01-MAR-1989 (Rel. 10, Created)

DT 01-FEB-1996 (Rel. 33, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Fibropellin I precursor (Epidermal growth factor-related protein 1) (UEGF-1).

GN EGF1.

OS Strongylocentrotus purpuratus (Purple sea urchin).

OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa; Echinoidea; Euechinoidea; Echinacea; Echinoda; Strongylocentrotidae;

OC Strongylocentrotus.

OX NCBI\_TaxID=7668;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90112459; PubMed=2514273;

RA Delgado-Illio-Reynoso M.G., Rollo D.R., Hursh D.A., Raff R.A.; "Structural analysis of the UEGF gene in the sea urchin Strongylocentrotus purpuratus reveals more similarity to vertebrate than to invertebrate genes with EGF-like repeats."

RT J. Mol. Evol. 29:314-327(1989).

RN [2]

RP SEQUENCE OF 279-476 AND 781-1064 FROM N.A.

RX MEDLINE=87319677; PubMed=3498216;

RA Hursh D.A., Andrews M.E., Raff R.A.; "A sea urchin gene encodes a polypeptide homologous to epidermal growth factor."

RT Science 237:1487-1490(1987).

RN [3]

RP AVIDIN-LIKE DOMAIN.

RX MEDLINE=89196806; PubMed=2784773;

RA Hunt L.T., Barker W.C.; "Avidin-like domain in an epidermal growth factor homolog from a sea urchin."

RT FASEB J. 3:1760-1764(1989).

RN [4]

RP CHARACTERIZATION.

RX MEDLINE=91285254; PubMed=2060714;

RA Bisgrove B.W., Andrews M.E., Raff R.A.;







```

FT DISULFID 453 462 BY SIMILARITY.
FT DISULFID 468 479 BY SIMILARITY.
FT DISULFID 473 488 BY SIMILARITY.
FT DISULFID 480 499 BY SIMILARITY.
FT DISULFID 505 515 BY SIMILARITY.
FT DISULFID 509 520 BY SIMILARITY.
FT DISULFID 522 531 BY SIMILARITY.
FT DISULFID 549 562 BY SIMILARITY.
FT DISULFID 556 569 BY SIMILARITY.
FT DISULFID 571 580 BY SIMILARITY.
FT DISULFID 586 597 BY SIMILARITY.
FT DISULFID 591 602 BY SIMILARITY.
FT DISULFID 604 610 BY SIMILARITY.
FT DISULFID 613 624 BY SIMILARITY.
FT DISULFID 618 634 BY SIMILARITY.
FT DISULFID 636 645 BY SIMILARITY.
FT DISULFID 652 664 BY SIMILARITY.
FT DISULFID 659 673 BY SIMILARITY.
FT DISULFID 675 684 BY SIMILARITY.
FT DISULFID 691 702 BY SIMILARITY.
FT DISULFID 696 711 BY SIMILARITY.
FT DISULFID 713 722 BY SIMILARITY.
FT DISULFID 729 740 BY SIMILARITY.
FT DISULFID 734 749 BY SIMILARITY.
FT DISULFID 751 760 BY SIMILARITY.
FT DISULFID 767 778 BY SIMILARITY.
FT DISULFID 772 787 BY SIMILARITY.
FT DISULFID 789 799 BY SIMILARITY.
FT DISULFID 806 817 BY SIMILARITY.
FT DISULFID 811 826 BY SIMILARITY.
FT DISULFID 828 837 BY SIMILARITY.
FT DISULFID 844 855 BY SIMILARITY.
FT DISULFID 849 880 BY SIMILARITY.
FT DISULFID 882 901 BY SIMILARITY.
FT DISULFID 908 919 BY SIMILARITY.
FT DISULFID 913 928 BY SIMILARITY.
FT DISULFID 930 939 BY SIMILARITY.
FT DISULFID 946 957 BY SIMILARITY.
FT DISULFID 952 966 BY SIMILARITY.
FT DISULFID 968 977 BY SIMILARITY.
FT DISULFID 984 995 BY SIMILARITY.
FT DISULFID 989 1009 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1211 1222 BY SIMILARITY.
FT DISULFID 1216 1231 BY SIMILARITY.
FT DISULFID 1233 1242 BY SIMILARITY.
FT DISULFID 1485 1496 BY SIMILARITY.
FT DISULFID 1490 1505 BY SIMILARITY.
FT DISULFID 1507 1516 BY SIMILARITY.
FT DISULFID 1574 1583 BY SIMILARITY.
FT DISULFID 1763 1774 BY SIMILARITY.
FT DISULFID 1768 1783 BY SIMILARITY.
FT DISULFID 1785 1794 BY SIMILARITY.
FT DISULFID 1801 1812 BY SIMILARITY.
FT DISULFID 1806 1821 BY SIMILARITY.
FT DISULFID 1833 1850 BY SIMILARITY.
FT DISULFID 1839 1859 BY SIMILARITY.
FT DISULFID 1844 1870 BY SIMILARITY.
FT DISULFID 1861 1889 BY SIMILARITY.
FT DISULFID 1878 1903 BY SIMILARITY.
FT DISULFID 1883 1914 BY SIMILARITY.
FT DISULFID 1905 1930 BY SIMILARITY.
FT DISULFID 1919 1939 BY SIMILARITY.
FT DISULFID 1924 1950 BY SIMILARITY.
FT DISULFID 1941 1968 BY SIMILARITY.
FT DISULFID 1957 1977 BY SIMILARITY.
FT DISULFID 1962 1988 BY SIMILARITY.
FT DISULFID 1979 2008 BY SIMILARITY.
FT DISULFID 1995 2017 BY SIMILARITY.
FT DISULFID 2002 2028 BY SIMILARITY.
FT DISULFID 2019 2037 BY SIMILARITY.
FT CARBOHYD 37 96 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 96 198 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 198 N-LINKED (GLCNAC. . .) (POTENTIAL).

```

```

FT CARBOHYD 238 238 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 239 239 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 336 336 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 400 400 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 40.0%; Score 86; DB 1; Length 2139;
Best Local Similarity 38.2%; Pred. No. 0.0052;
Matches 13; Conservative 6; Mismatches 15; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCEIINNHTCNDGVYGPQCL 34
Db 908 CDSNPSCKHNCNDIGCTYTCCEPGEPTHCEI 941

RESULT 29
NOTC_DROME STANDARD; PRT; 2703 AA.
AC P07207; P04154; 097458; 09W4T8;
DT 01-NOV-1986 (Rel. 03, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Neurogenic locus Notch protein precursor.
GN N OR EG:140G11.1 OR EG:163A10.2 OR CG3936.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_Taxid=7227;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN-Oregon-R; TISSUE=Embryo;
RX MEDLINE=86079359; PubMed=3935325;
RA Wharton K.A., Johansen K.M., Xu T., Artavanis-Tsakonas S.;
RT "Nucleotide sequence from the neurogenic locus notch implies a gene
RT product that shares homology with proteins containing EGF-like
RT repeats."
RN Cell 43:567-581(1985).
[2]
RP SEQUENCE FROM N.A.
RC STRAIN-Canton-S; and Oregon-R; TISSUE=Embryo;
RX MEDLINE=87064624; PubMed=3097517;
RA Kidd S., Kelley M.R., Young M.W.;
RT "Sequence of the notch locus of Drosophila melanogaster: relationship
RT of the encoded protein to mammalian clotting and growth factors."
RN Mol. Cell. Biol. 6:3094-3108(1986).
[3]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.A., Butler J., Brokstein P., Brotlier P.,
RA Burris K.C., Busch D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jaitani M., Kalush F., Kapen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodali C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

```



Best Local Similarity 42.4%; Pred. No. 0.0064;  
Matches 14; Conservative 4; Mismatches 15; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGEVLEIINHTCNCVGYGPOCQ 33

Db 1026 CRESSCNGSCIDINGNCISLAGSGANCO 1058

RESULT 30

FP2\_MYTGA STANDARD: PRT: 473 AA.

ID FP2\_MYTGA

AC Q25464;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Adhesive plaque matrix protein 2 precursor (Foot protein 2) (MGFP2)

DE (MGFP-2).

GN FP2.

OS Mytilus galloprovincialis (Mediterranean mussel).

OC Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorpha; Mytiloidea;

OC Mytiloidea; Mytilidae; Mytilus.

OK NCBI\_TaxID=29158;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Foot;

RX MEDLINE=95204464; PubMed=7896812;

RA Inoue K., Takeuchi Y., Miki D., Odo S.;

"Mussel adhesive plaque protein gene is a novel member of epidermal

growth factor-like gene family.";

RL J. Biol. Chem. 270:6698-6701(1995).

-1 FUNCTION: PROVIDES ADHESIVENESS TO THE MUSSEL'S FOOT. MUSSELS

PRODUCE ONE OF THE STRONGEST WATER INSOLUBLE GLUE. THE MUSSEL'S

ADHESIVE IS A BUNDLE OF THREADS, CALLED A BYSSUS, FORMED BY A

FIBROUS COLLAGENOUS CORE COATED WITH ADHESIVE PROTEINS.

-1 SUBCELLULAR LOCATION: Secreted.

-1 TISSUE SPECIFICITY: PRODUCED BY THE BYSSAL GLAND.

-1 DEVELOPMENTAL STAGE: EXPRESSION STARTS AT THE PEDIVELIGER, FOOT

FORMATION, STAGE.

-1 PTM: SOME TYROSINES ARE HYDROXYLATED (THUS PRODUCING DOPA - 3,4-

DIHYDROXYPHENYLALANINE).

-1 SIMILARITY: CONTAINS 11 EGF-LIKE DOMAINS.

-----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: DA3794; BAA07852.1; -.

DR HSSP: P00740; IEDM.

DR InterPro: IPR000152; Asx\_hydroxyl.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR000742; EGF-2.

DR InterPro: IPR001881; EGF\_Ca.

DR Pfam: PF00008; EGF\_11.

DR SMART: SM00181; EGF\_10.

DR PROSITE: PS00010; Asx\_hydroxyl; 2.

DR PROSITE: PS00022; EGF\_1; 11.

DR PROSITE: PS01186; EGF\_2; 10.

KW Glycoprotein; EGF-like domain; Repeat; Signal; Hydroxylation.

FT SIGNAL 1 17 POTENTIAL.

FT CHAIN 18 473 ADHESIVE PLAUQUE MATRIX PROTEIN 2.

FT DOMAIN 45 81 EGF-LIKE 1.

FT DOMAIN 82 117 EGF-LIKE 2.

FT DOMAIN 118 154 EGF-LIKE 3.

FT DOMAIN 155 191 EGF-LIKE 4.

FT DOMAIN 192 228 EGF-LIKE 5.

FT DOMAIN 229 265 EGF-LIKE 6.

FT DOMAIN 266 301 EGF-LIKE 7.

FT DOMAIN 302 340 EGF-LIKE 8.

FT DOMAIN 342 378 EGF-LIKE 9.

FT DOMAIN 383 420 EGF-LIKE 10.

FT DOMAIN 425 461 EGF-LIKE 11.

FT MOD\_RES 23 DOPA.

FT MOD\_RES 31 DOPA.

FT MOD\_RES 36 DOPA.

FT MOD\_RES 43 DOPA.

FT CARBOHYD 93 93 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT DISULFID 49 60 BY SIMILARITY.

FT DISULFID 54 69 BY SIMILARITY.

FT DISULFID 71 80 BY SIMILARITY.

FT DISULFID 86 97 BY SIMILARITY.

FT DISULFID 91 106 BY SIMILARITY.

FT DISULFID 108 117 BY SIMILARITY.

FT DISULFID 122 133 BY SIMILARITY.

FT DISULFID 127 143 BY SIMILARITY.

FT DISULFID 145 154 BY SIMILARITY.

FT DISULFID 159 170 BY SIMILARITY.

FT DISULFID 164 180 BY SIMILARITY.

FT DISULFID 182 191 BY SIMILARITY.

FT DISULFID 196 207 BY SIMILARITY.

FT DISULFID 201 217 BY SIMILARITY.

FT DISULFID 219 228 BY SIMILARITY.

FT DISULFID 233 244 BY SIMILARITY.

FT DISULFID 238 254 BY SIMILARITY.

FT DISULFID 256 265 BY SIMILARITY.

FT DISULFID 270 281 BY SIMILARITY.

FT DISULFID 275 290 BY SIMILARITY.

FT DISULFID 292 301 BY SIMILARITY.

FT DISULFID 306 317 BY SIMILARITY.

FT DISULFID 311 328 BY SIMILARITY.

FT DISULFID 330 339 BY SIMILARITY.

FT DISULFID 346 357 BY SIMILARITY.

FT DISULFID 351 366 BY SIMILARITY.

FT DISULFID 368 377 BY SIMILARITY.

FT DISULFID 387 399 BY SIMILARITY.

FT DISULFID 393 408 BY SIMILARITY.

FT DISULFID 410 419 BY SIMILARITY.

FT DISULFID 429 440 BY SIMILARITY.

FT DISULFID 434 449 BY SIMILARITY.

FT DISULFID 451 460 BY SIMILARITY.

SO SEQUENCE 473 AA; 51772 MW; BA76BA8C3BA9A0F CRC64;

Query Match 39.8%; Score 85.5; DB 1; Length 473;

Best Local Similarity 38.2%; Pred. No. 0.0015;

Matches 13; Conservative 8; Mismatches 12; Indels 1; Gaps 1;

OY 1 CQPMSCSGHGEVLEIINHTCNCVGYGPOCQ 33

Db 196 CKPNCNSNGTCSADKFDYSCENRPTGPECE 229

RESULT 31

TENA\_HUMAN

ID TENA\_HUMAN

AC P24821; Q15567; Q14583;

DT 01-MAR-1992 (Rel. 21, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Tenascin precursor (TN) (Hexabrachion) (Cytotactin) (Neurexectin)

DE (GWM) (Jf) (Microendous antigen) (Glioma-associated-extracellular

matrix antigen) (GP 150-225) (Tenascin-C) (TN-C).

GN HXB.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OK NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-32.

RC TISSUE=Fetal brain, and Melanoma;

RX MEDLINE=91187670; PubMed=1707164;

RA Sirtl A., Carmelita B., Saginatti M., Leprini A., Casari G.,

Baralle F., Zardi L.;



```
FT DISULFID 456 465 BY SIMILARITY.
FT DISULFID 470 480 BY SIMILARITY.
FT DISULFID 474 485 BY SIMILARITY.
FT DISULFID 487 496 BY SIMILARITY.
FT DISULFID 501 511 BY SIMILARITY.
FT DISULFID 505 516 BY SIMILARITY.
FT DISULFID 518 527 BY SIMILARITY.
FT DISULFID 532 542 BY SIMILARITY.
FT DISULFID 536 547 BY SIMILARITY.
FT DISULFID 549 558 BY SIMILARITY.
FT DISULFID 563 573 BY SIMILARITY.
FT DISULFID 567 578 BY SIMILARITY.
FT DISULFID 580 589 BY SIMILARITY.
FT DISULFID 594 604 BY SIMILARITY.
FT DISULFID 598 609 BY SIMILARITY.
FT DISULFID 611 620 BY SIMILARITY.
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 166 166 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 184 184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 327 327 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 788 788 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1018 1018 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1034 1034 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1079 1079 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1093 1093 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1119 1119 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1184 1184 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1210 1210 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1261 1261 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1275 1275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1301 1301 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1366 1366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1392 1392 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1445 1445 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1455 1455 N-LINKED (GLCNAC. . .) (POTENTIAL).
```

Query Match 39.1%; Score 84; DB 1; Length 2201;  
Best Local Similarity 53.3%; Pred. No. 0.0094;  
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;

```
Qy 3 PWSGSGHGCEVETINHTGNCDDVGYGPOC 32
Db 409 PNGSGHGRCV---NGQCVDEGYTGEDC 434
```

```
RESULT 32
NTC3_MOUSE STANDARD: PRT: 2318 AA.
ID 061982;
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Neurogenic locus notch 3 protein.
GN NOTCH3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ICR X SWISS WEBSTER;
RX MEDLINE=95001556; PubMed=7918097;
RA Larelli M., Dalstrand J., Lendahl U.;
RT "The novel Notch homologue mouse Notch 3 lacks specific epidermal
RT growth factor-repeats and is expressed in proliferating
RT neuroepithelium."
RL Mech. Dev. 46:123-136(1994).
CC -1- FUNCTION: NOTCH 1, 2 AND 3 PLAY A COMBINATIONAL ROLE DURING
CC VARIOUS CELL FATE DECISIONS AND MORPHOLOGICAL MOVEMENTS IN THE
CC DEVELOPING CNS AND PROBABLY OTHER REGIONS OF THE EMBRYO.
CC -1- TISSUE SPECIFICITY: PROLIFERATING NEUROEPITHELIUM.
CC -1- DEVELOPMENTAL STAGE: CNS DEVELOPMENT.
CC -1- SIMILARITY: CONTAINS 34 EGF-LIKE DOMAINS.
```

```
CC -1- SIMILARITY: CONTAINS 3 LIN/NOTCH REPEATS.
CC -1- SIMILARITY: CONTAINS 5 ANK REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
CC EMBL: X74760; CAA52776.1; -.
CC HSSP: P00740; 11XA.
CC MGD: MGI:99460; Notch3.
DR InterPro: IPR002110; ANK.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR000742; EGF-2.
DR InterPro: IPR001881; EGF_Ca.
DR InterPro: IPR001438; EGF-II.
DR InterPro: IPR000800; Notch.
DR Pfam: PF00023; ank; 6.
DR Pfam: PF00008; EGF; 34.
DR Pfam: PF00066; notch; 3.
DR PRINTS: PR00010; EGFLOOD.
DR PRINTS: PR01452; NOTCH.
DR SMART: SM00248; ANK; 5.
DR SMART: SM00179; EGF_CA; 19.
DR SMART: SM00001; EGF-like; 13.
DR SMART: SM00004; NL; 3.
DR PROSITE: PS50088; ANK_REPEAT; 4.
DR PROSITE: PS50297; ANK_REPEAT_REGION; 1.
DR PROSITE: PS00022; EGF-1; 33.
DR PROSITE: PS01186; EGF-2; 27.
DR PROSITE: PS01187; EGF_CA; 17.
DR PROSITE: PS01187; Neurogenesis; Repeat; EGF-like domain; Transmembrane;
KW ANK repeat; Glycoprotein.
FT DOMAIN 1 1643 EXTRACELLULAR.
FT TRANSMEM 1644 1664 POTENTIAL.
FT DOMAIN 1665 2318 CYTOPLASMIC.
FT DOMAIN 2242 2261 PEST.
FT DOMAIN 39 78 EGF-LIKE 1.
FT DOMAIN 79 119 EGF-LIKE 2.
FT DOMAIN 120 157 EGF-LIKE 3.
FT DOMAIN 159 196 EGF-LIKE 4.
FT DOMAIN 235 295 EGF-LIKE 5.
FT DOMAIN 237 273 EGF-LIKE 6.
FT DOMAIN 275 313 EGF-LIKE 7.
FT DOMAIN 315 351 EGF-LIKE 8.
FT DOMAIN 352 380 EGF-LIKE 9.
FT DOMAIN 392 430 EGF-LIKE 10.
FT DOMAIN 432 468 EGF-LIKE 11.
FT DOMAIN 470 506 EGF-LIKE 12.
FT DOMAIN 508 544 EGF-LIKE 13.
FT DOMAIN 546 581 EGF-LIKE 14.
FT DOMAIN 583 619 EGF-LIKE 15.
FT DOMAIN 621 656 EGF-LIKE 16.
FT DOMAIN 658 694 EGF-LIKE 17.
FT DOMAIN 696 731 EGF-LIKE 18.
FT DOMAIN 735 771 EGF-LIKE 19.
FT DOMAIN 772 809 EGF-LIKE 20.
FT DOMAIN 811 848 EGF-LIKE 21.
FT DOMAIN 850 886 EGF-LIKE 22.
FT DOMAIN 923 961 EGF-LIKE 23.
FT DOMAIN 963 999 EGF-LIKE 24.
FT DOMAIN 1001 1035 EGF-LIKE 25.
FT DOMAIN 1037 1083 EGF-LIKE 26.
FT DOMAIN 1085 1121 EGF-LIKE 27.
FT DOMAIN 1123 1159 EGF-LIKE 28.
FT DOMAIN 1161 1204 EGF-LIKE 29.
FT DOMAIN 1206 1245 EGF-LIKE 30.
FT DOMAIN 1245 EGF-LIKE 31.
```

FT DOMAIN 1247 1288 EGF-LIKE 32.  
FT DOMAIN 1290 1326 EGF-LIKE 33.  
FT DOMAIN 1336 1374 EGF-LIKE 34.  
FT REPEAT 1388 1428 LIN/NOTCH 1.  
FT REPEAT 1429 1467 LIN/NOTCH 2.  
FT REPEAT 1468 1503 LIN/NOTCH 3.  
FT REPEAT 1839 1868 ANK 1.  
FT REPEAT 1872 1902 ANK 2.  
FT REPEAT 1906 1935 ANK 3.  
FT REPEAT 1939 1968 ANK 4.  
FT REPEAT 1972 2001 ANK 5.  
FT DISULFID 43 55 BY SIMILARITY.  
FT DISULFID 49 66 BY SIMILARITY.  
FT DISULFID 63 77 BY SIMILARITY.  
FT DISULFID 88 94 BY SIMILARITY.  
FT DISULFID 109 118 BY SIMILARITY.  
FT DISULFID 124 135 BY SIMILARITY.  
FT DISULFID 129 145 BY SIMILARITY.  
FT DISULFID 147 156 BY SIMILARITY.  
FT DISULFID 163 175 BY SIMILARITY.  
FT DISULFID 169 184 BY SIMILARITY.  
FT DISULFID 186 195 BY SIMILARITY.  
FT DISULFID 202 213 BY SIMILARITY.  
FT DISULFID 207 223 BY SIMILARITY.  
FT DISULFID 225 234 BY SIMILARITY.  
FT DISULFID 241 252 BY SIMILARITY.  
FT DISULFID 246 261 BY SIMILARITY.  
FT DISULFID 263 272 BY SIMILARITY.  
FT DISULFID 279 292 BY SIMILARITY.  
FT DISULFID 286 301 BY SIMILARITY.  
FT DISULFID 303 312 BY SIMILARITY.  
FT DISULFID 319 330 BY SIMILARITY.  
FT DISULFID 324 339 BY SIMILARITY.  
FT DISULFID 341 350 BY SIMILARITY.  
FT DISULFID 356 367 BY SIMILARITY.  
FT DISULFID 361 378 BY SIMILARITY.  
FT DISULFID 380 389 BY SIMILARITY.  
FT DISULFID 396 409 BY SIMILARITY.  
FT DISULFID 403 418 BY SIMILARITY.  
FT DISULFID 420 429 BY SIMILARITY.  
FT DISULFID 436 447 BY SIMILARITY.  
FT DISULFID 441 456 BY SIMILARITY.  
FT DISULFID 458 467 BY SIMILARITY.  
FT DISULFID 474 485 BY SIMILARITY.  
FT DISULFID 479 494 BY SIMILARITY.  
FT DISULFID 496 505 BY SIMILARITY.  
FT DISULFID 512 523 BY SIMILARITY.  
FT DISULFID 517 532 BY SIMILARITY.  
FT DISULFID 534 543 BY SIMILARITY.  
FT DISULFID 550 560 BY SIMILARITY.  
FT DISULFID 555 569 BY SIMILARITY.  
FT DISULFID 571 580 BY SIMILARITY.  
FT DISULFID 587 598 BY SIMILARITY.  
FT DISULFID 592 607 BY SIMILARITY.  
FT DISULFID 609 618 BY SIMILARITY.  
FT DISULFID 625 635 BY SIMILARITY.  
FT DISULFID 630 644 BY SIMILARITY.  
FT DISULFID 646 655 BY SIMILARITY.  
FT DISULFID 662 673 BY SIMILARITY.  
FT DISULFID 667 682 BY SIMILARITY.  
FT DISULFID 684 693 BY SIMILARITY.  
FT DISULFID 700 710 BY SIMILARITY.  
FT DISULFID 705 719 BY SIMILARITY.  
FT DISULFID 721 730 BY SIMILARITY.  
FT DISULFID 739 750 BY SIMILARITY.  
FT DISULFID 744 759 BY SIMILARITY.  
FT DISULFID 761 770 BY SIMILARITY.  
FT DISULFID 776 787 BY SIMILARITY.  
FT DISULFID 781 797 BY SIMILARITY.  
FT DISULFID 799 808 BY SIMILARITY.  
FT DISULFID 815 827 BY SIMILARITY.  
FT DISULFID 821 836 BY SIMILARITY.

FT DISULFID 838 847 BY SIMILARITY.  
FT DISULFID 854 865 BY SIMILARITY.  
FT DISULFID 859 874 BY SIMILARITY.  
FT DISULFID 876 885 BY SIMILARITY.  
FT DISULFID 892 902 BY SIMILARITY.  
FT DISULFID 897 911 BY SIMILARITY.  
FT DISULFID 913 922 BY SIMILARITY.  
FT DISULFID 929 940 BY SIMILARITY.  
FT DISULFID 934 949 BY SIMILARITY.  
FT DISULFID 951 960 BY SIMILARITY.  
FT DISULFID 967 978 BY SIMILARITY.  
FT DISULFID 972 987 BY SIMILARITY.  
FT DISULFID 989 998 BY SIMILARITY.  
FT DISULFID 1005 1016 BY SIMILARITY.  
FT DISULFID 1010 1023 BY SIMILARITY.  
FT DISULFID 1025 1034 BY SIMILARITY.  
FT DISULFID 1041 1062 BY SIMILARITY.  
FT DISULFID 1056 1071 BY SIMILARITY.  
FT DISULFID 1073 1082 BY SIMILARITY.  
FT DISULFID 1089 1100 BY SIMILARITY.  
FT DISULFID 1094 1109 BY SIMILARITY.  
FT DISULFID 1111 1120 BY SIMILARITY.  
FT DISULFID 1127 1138 BY SIMILARITY.  
FT DISULFID 1132 1147 BY SIMILARITY.  
FT DISULFID 1149 1158 BY SIMILARITY.  
FT DISULFID 1165 1183 BY SIMILARITY.  
FT DISULFID 1177 1192 BY SIMILARITY.  
FT DISULFID 1194 1203 BY SIMILARITY.

Query Match 39.1%; Score 84; DB 1; Length 2318;  
Best Local Similarity 42.4%; Pred. No. 0.0099;  
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGCVETIINNNHTCNCDVGYGPQCQ 33  
DB 929 CSPSSCFNGGCVDDVSSFCLOCPGYGTGHQ 961

RESULT 33  
ID PRTC\_RAT STANDARD; PRT; 461 AA.  
AC P31394;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Vitamin-K dependent protein C precursor (EC 3.4.21.69)  
DE (Autoproteolysin IIA) (Anticoagulant protein C) (Blood coagulation factor XIV).  
DE PROC.  
GN PROC.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-WISTAR; TISSUE=Liver;  
RX MEDLINE=92329550; PubMed=1627650;  
RA Okafuji T., Maekawa K., Nawa K., Marumoto Y.;  
RT "The CDNA cloning and mRNA expression of rat protein C.";  
RU Biochim. Biophys. Acta 1131:329-332(1992).  
CC -1- FUNCTION: PROTEIN C IS A VITAMIN K-DEPENDENT SERINE PROTEASE THAT  
CC REGULATES BLOOD COAGULATION BY INACTIVATING FACTORS VA AND VIIIA  
CC IN THE PRESENCE OF CALCIUM IONS AND PHOSPHOLIPIDS.  
CC -1- CATALYTIC ACTIVITY: Degradation of blood coagulation factors Va  
CC and VIIIA.  
CC -1- SUBUNIT: SYNTHESIZED AS A SINGLE CHAIN PRECURSOR, WHICH IS CLEAVED  
CC INTO A LIGHT CHAIN AND A HEAVY CHAIN HELD TOGETHER BY A DISULFIDE  
CC BOND. THE ENZYME IS THEN ACTIVATED BY THROMBIN, WHICH CLEAVES A  
CC TETRADECAPEPTIDE FROM THE AMINO END OF THE HEAVY CHAIN; THIS  
CC REACTION, WHICH OCCURS AT THE SURFACE OF ENDOTHELIAL CELLS, IS  
CC STRONGLY PROMOTED BY THROMBOMODULIN.  
CC -1- TISSUE SPECIFICITY: PLASMA; SYNTHESIZED IN THE LIVER.  
CC -1- PTM: THE VITAMIN K-DEPENDENT, ENZYMIC CARBOXYLATION OF SOME

CC GLU RESIDUES ALLOWS THE MODIFIED PROTEIN TO BIND CALCIUM.  
 CC -1- MISCELLANEOUS: CALCIUM ALSO BINDS, WITH STRONGER AFFINITY TO  
 CC ANOTHER SITE, BEYOND THE GLA DOMAIN. THIS GLA-INDEPENDENT BINDING  
 CC SITE IS NECESSARY FOR THE RECOGNITION OF THE  
 CC THROMBIN-THROMBOMODULIN COMPLEX.  
 CC -1- SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
 CC TRYPSIN FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X64336; CAA45617.1; -.  
 CC PIR: S18994; S18994.  
 CC PIR: S24312; S24312.  
 CC HSSP: P04070; LPCU.  
 CC MEROPS: S01\_218; -.  
 CC InterPro: IPR000152; Asx\_hydroxyl.  
 CC InterPro: IPR001314; Chymotrypsin.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR001881; EGF\_Ca.  
 CC InterPro: IPR002383; GLA-blood.  
 CC InterPro: IPR001254; Trypsin.  
 CC InterPro: IPR000294; Vitk\_dep\_GLA.  
 CC Pfam: PF00008; EGF; 2.  
 CC Pfam: PF00594; gla; 1.  
 CC Pfam: PF00089; trypsin; 1.  
 CC PRINTS: PR00722; CHYMOTRYPSIN.  
 CC PRINTS: PR00001; GLABLOOD.  
 CC SMART: SM00179; EGF\_CA; 1.  
 CC SMART: SM00001; EGF-like; 1.  
 CC SMART: SM00069; GLA; 1.  
 CC SMART: SM00020; TRYP\_SPE; 1.  
 CC PROSITE: PS00010; ASX\_HYDROXYL; 1.  
 CC PROSITE: PS00022; EGF\_1; 1.  
 CC PROSITE: PS01186; EGF\_2; 2.  
 CC PROSITE: PS01187; EGF\_CA; 1.  
 CC PROSITE: PS00011; GLU\_CARBOXYLATION; 1.  
 CC PROSITE: PS00240; TRYPSIN\_DOM; 1.  
 CC PROSITE: PS00134; TRYPSIN\_HIS; 1.  
 CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
 CC Blood coagulation; Glycoprotein; Serine protease;  
 CC Gamma-carboxylglutamic acid; Calcium-binding; Vitamin K; Hydroxylation;  
 CC EGF-like domain; Repeat; Endothelial cell; Hyaluronase; Signal.  
 CC EGF-like domain; Repeat; Endothelial cell; Hyaluronase; Signal.  
 CC PROPEP 1 33 41  
 CC CHAIN 42 196  
 CC CHAIN 199 461  
 CC PEPTIDE 199 212  
 CC SITE 212 213  
 CC DOMAIN 96 131  
 CC DOMAIN 135 175  
 CC DOMAIN 213 461  
 CC MOD\_RES 47 47  
 CC MOD\_RES 48 48  
 CC MOD\_RES 55 55  
 CC MOD\_RES 57 57  
 CC MOD\_RES 60 60  
 CC MOD\_RES 61 61  
 CC MOD\_RES 66 66  
 CC MOD\_RES 67 67

FT MOD\_RES 70 70 (BY SIMILARITY).  
 FT ACT\_SITE 112 112 GAMMA-CARBOXYGLUTAMIC ACID  
 FT ACT\_SITE 254 254 (BY SIMILARITY).  
 FT ACT\_SITE 300 300 HYDROXYLATION (BY SIMILARITY).  
 FT ACT\_SITE 402 402 CHARGE RELAY SYSTEM.  
 FT DISULFID 58 63 CHARGE RELAY SYSTEM.  
 FT DISULFID 91 110 BY SIMILARITY.  
 FT DISULFID 100 105 BY SIMILARITY.  
 FT DISULFID 104 119 BY SIMILARITY.  
 FT DISULFID 121 130 BY SIMILARITY.  
 FT DISULFID 139 150 BY SIMILARITY.  
 FT DISULFID 146 159 BY SIMILARITY.  
 FT DISULFID 161 174 BY SIMILARITY.  
 FT DISULFID 182 320 INTERCHAIN (BY SIMILARITY).  
 FT DISULFID 239 255 BY SIMILARITY.  
 FT DISULFID 373 387 BY SIMILARITY.  
 FT DISULFID 398 426 BY SIMILARITY.  
 FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 291 291 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 355 355 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 461 AA; 51912 MW; 8A4CF9364EDACD5 CRC64;  
 Query Match 38.6%; Score 83; DB 1; Length 461;  
 Best Local Similarity 42.9%; Pred No. 0.003; Mismatches 10; Indels 0; Gaps 0;  
 Matches 12; Conservative 6; Mismatches 10; Indels 0; Gaps 0;  
 QY 6 CSGHGCEVEIINHTNCNDVGYGPOCQ 33  
 DB 104 CCGHGTCDIGLGGFSCGCKGMEGRQC 131  
 ID FBP3\_STRPU STANDARD; PRT; 570 AA.  
 AC P49013;  
 DT 01-FEB-1996 (rel. 33, last sequence update)  
 DT 16-OCT-2001 (rel. 40, last annotation update)  
 DE Fibropellin C precursor (Epidermal growth factor-related protein 3)  
 DE (EGF III) (Fibropellin III).  
 GN EGF3.  
 OS Strongylocentrotus purpuratus (Purple sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eutherozoa; Echinozoa;  
 OC Echinoidea; Echinoidea; Echinacea; Echinoidea; Strongylocentrotidae;  
 OC Strongylocentrotus.  
 OX NCBI\_TaxID=7668;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Gastrula;  
 RA MEDLINE=93273088; PubMed=8500658;  
 RX Bisgrove B.W., Rafi R.A.;  
 RT "The SPEGF III gene encodes a member of the fibropellins: EGF repeat-  
 RT containing proteins that form the apical lamina of the sea urchin  
 RT embryo.";  
 RL Dev. Biol. 157:526-538(1993).  
 CC -1- FUNCTION: FORM THE APICAL LAMINA, A COMPONENT OF THE EXTRACELLULAR  
 CC MATRIX.  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- DEVELOPMENTAL STAGE: LOW LEVELS IN UNFERTILIZED EGGS AND DURING  
 CC EARLY CLEAVAGE, THEN RAPIDLY INCREASES IN ABUNDANCE BETWEEN LATE  
 CC MORULA AND MESENCHYME BLASTULA STAGES TO MAXIMAL LEVELS MAINTAINED  
 CC THROUGH SUBSEQUENT STAGES.  
 CC -1- MISCELLANEOUS: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY.  
 CC -1- SIMILARITY: CONTAINS 8 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 1 CUB DOMAIN.  
 CC -1- SIMILARITY: THE C-TERMINAL DOMAIN OF THIS PROTEIN IS SIMILAR  
 CC TO AVIDIN/STREPTAVIDIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: L07045; AAA30045.1; -.  
 DR HSBP; P00740; IEDM.  
 DR InterPro: IPR000152; Asx\_hydroxyl.  
 DR InterPro: IPR000088; Avidin.  
 DR InterPro: IPR000859; CUB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF-2.  
 DR InterPro: IPR001881; EGF\_Ca.  
 DR InterPro: IPR001438; EGF\_II.  
 DR Pfam: PF01382; Avidin; 1.  
 DR Pfam: PF00431; CUB; 1.  
 DR Pfam: PF00008; EGF; 8.  
 DR PRINTS; PR00010; EGFBLDOD.  
 DR ProDom: PD016055; Avidin; 1.  
 DR SMART; SM00042; CUB; 1.  
 DR SMART; SM00179; EGF\_Ca; 7.  
 DR SMART; SM00001; EGF-like; 1.  
 DR PROSITE; PS00022; EGF\_1; 8.  
 DR PROSITE; PS00577; AVIDIN; 1.  
 DR PROSITE; PS01180; CUB; 1.  
 DR PROSITE; PS01186; EGF-2; 7.  
 DR PROSITE; PS01187; EGF\_Ca; 6.  
 DR Biotin: EGF-like domain; Repeat; Signal; Glycoprotein;  
 KW Calcium-binding.  
 FT SIGNAL 1 17  
 FT CHAIN 18 570  
 FT DOMAIN 18 55  
 FT DOMAIN 62 175  
 FT DOMAIN 176 212  
 FT DOMAIN 214 250  
 FT DOMAIN 252 288  
 FT DOMAIN 290 326  
 FT DOMAIN 328 364  
 FT DOMAIN 366 402  
 FT DOMAIN 404 440  
 FT DOMAIN 442 570  
 FT DISULFID 23 34  
 FT DISULFID 28 43  
 FT DISULFID 45 54  
 FT DISULFID 180 191  
 FT DISULFID 185 200  
 FT DISULFID 202 211  
 FT DISULFID 216 229  
 FT DISULFID 223 238  
 FT DISULFID 240 249  
 FT DISULFID 256 267  
 FT DISULFID 261 276  
 FT DISULFID 278 287  
 FT DISULFID 294 305  
 FT DISULFID 299 314  
 FT DISULFID 316 325  
 FT DISULFID 332 343  
 FT DISULFID 337 352  
 FT DISULFID 354 363  
 FT DISULFID 370 381  
 FT DISULFID 375 390  
 FT DISULFID 392 401  
 FT DISULFID 408 419  
 FT DISULFID 413 428  
 FT DISULFID 430 439  
 FT CARBOHYD 30 30  
 FT CARBOHYD 136 136  
 FT CARBOHYD 357 357  
 SQ SEQUENCE 570 AA; 61116 MW; BE653E3E1C05E6EE CRC64;

Query Match 38.6%; Score 83; DB 1; Length 570;  
 Best Local Similarity 39.4%; Pred. No. 0.0036;  
 Matches 13; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

OY 1 CPMSCSGHECVETIINHTNCNDVGYGPOCQ 33

Db 218 CASAPCRNGACVDQVNGVTCNCPFNGVNC 250

RESULT 35  
 ID CRBH\_HUMAN STANDARD; PRT; 1376 AA.  
 AC P82279;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Crumbs protein homolog 1 precursor.  
 GN CRBL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND VARIANTS RP12 V-161; W-250; M-745; C-764;  
 RP Y-948; T-1041 AND P-1071.  
 RC TISSUE=Retina, and Fetal brain;  
 RX MEDLINE=99438399; PubMed=10508521;  
 RA den Hollander A.I., ten Brink J.B., de Kok Y.J.M., van Soest S.,  
 RA van den Born L.I., van Driel M.A., van de Pol D.J.R., Payne A.M.,  
 RA Bhattacharya S.S., Kellner U., Hoyng C.B., Westerveld A.,  
 RA Brunner H.G., Bleeker-Wagemakers E.M., Deutman A.F.,  
 RA Heckenlied J.R., Cremers F.P.M., Bergen A.B.;  
 RA "Mutations in a human homologue of drosophila crumbs cause retinitis  
 RA pigmentosa (Rp12)";  
 RT Nat. Genet. 23:217-221(1999).  
 RL [2]  
 RN VARIANT LCA ARG-1100, AND VARIANT RP ARG-1181.  
 RP MEDLINE=21303018; PubMed=11389483;  
 RX den Hollander A.I., Heckenlied J.R., van den Born L.I.,  
 RA de Kok Y.J.M., van der Velde-Visser S.D., Kellner U.,  
 RA van Schooneveld M.J., Blankenagel A., Kohnscheider K., Wissinger B.,  
 RA Cruysberg J.R.M., Deutman A.F., Brunner H.G., Apfelstedt-Sylla E.,  
 RA Hoyng C.B., Cremers F.P.M.;  
 RA "Leber congenital amaurosis and retinitis pigmentosa with Coats-like  
 RA exudative vasculopathy are associated with mutations in the crumbs  
 RL homologue 1 (CRBL) gene";  
 RT Am. J. Hum. Genet. 69:198-203(2001).  
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL-CELL INTERACTION IN NEURONAL  
 CC DEVELOPMENT OF THE RETINA.  
 CC -1- SUBCELLULAR LOCATION: Extracellular (Potential).  
 CC -1- TISSUE SPECIFICITY: PREFERENTIAL EXPRESSION IN RETINA, ALSO  
 CC EXPRESSED IN BRAIN AND FETAL BRAIN.  
 CC -1- DISEASE: DEFECTS IN CRBL ARE THE CAUSE OF RETINITIS PIGMENTOSA  
 CC TYPE 12 (RP12), AN AUTOSOMAL RECESSIVE CONDITION WHICH IS  
 CC CHARACTERIZED BY NIGHT BLINDNESS FROM EARLY CHILDHOOD AND  
 CC PROGRESSIVE VISUAL FIELD LOSS. THERE IS GENERAL LOSS OF RETINAL  
 CC PIGMENT EPITHELIUM THROUGHOUT THE RETINA AND PATIENTS EXPERIENCE  
 CC SEVERE VISUAL IMPAIRMENT BEFORE THE AGE OF TWENTY.  
 CC -1- DISEASE: Defects in CRBL are a cause of a form of Leber congenital  
 CC amaurosis (LCA).  
 CC -1- SIMILARITY: CONTAINS 15 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 LAMININ G-LIKE DOMAINS.  
 CC -1- DATABASE: NAME-Mutations of the CRBL gene;  
 CC NOTE=Retina International's Scientific Newsletter;  
 CC WWW="http://www.retina-international.com/sci-news/crblmut.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC -----  
DR EMBL: AF154671; AA01361.1; -  
DR HSSP: P08709; 1BP9.  
DR MIM: 604210; -  
DR MIM: 600105; -  
DR InterPro: IPR000152; Asx\_hydroxyl.  
DR InterPro: IPR0000561; EGF-like.  
DR InterPro: IPR000742; EGF\_2.  
DR InterPro: IPR001881; EGF\_Ca.  
DR InterPro: IPR001438; EGF\_III.  
DR InterPro: IPR002049; Laminin\_EGF.  
DR InterPro: IPR001791; Laminin\_G.  
DR Pfam: PF00008; EGF; 16.  
DR Pfam: PF00054; Laminin\_G; 3.  
DR PRINTS: PR00010; EGFBLD.  
DR PRINTS: PR00011; EGF\_LAMININ.  
DR SMART: SM00179; EGF\_CA; 8.  
DR SMART: SM00001; EGF\_Like; 8.  
DR SMART: SM00282; LamG; 3.  
DR PROSITE: PS00010; ASX\_HYDROXYL; 10.  
DR PROSITE: PS00022; EGF\_1; 15.  
DR PROSITE: PS01186; EGF\_2; 11.  
DR PROSITE: PS01187; EGF\_CA; 7.  
DR PROSITE: PS50025; Lam\_G\_DOMAIN; 3.  
KW EGF-like domain; Glycoprotein; Repeat; Signal; Disease mutation;  
KW Retinitis pigmentosa; Vision.  
FT SIGNAL 1 25 POTENTIAL.  
FT CHAIN 26 1376 CRUMBS PROTEIN HOMOLOG 1.  
FT DOMAIN 70 108 EGF-Like 1.  
FT DOMAIN 110 146 EGF-Like 2.  
FT DOMAIN 148 184 EGF-Like 3.  
FT DOMAIN 186 222 EGF-Like 4.  
FT DOMAIN 224 260 EGF-Like 5.  
FT DOMAIN 304 337 EGF-Like 6.  
FT DOMAIN 339 395 EGF-Like 7.  
FT DOMAIN 441 481 EGF-Like 8.  
FT DOMAIN 485 670 LAMININ G-Like 1.  
FT DOMAIN 672 708 EGF-Like 9.  
FT DOMAIN 714 885 LAMININ G-Like 2.  
FT DOMAIN 887 923 EGF-Like 10.  
FT DOMAIN 950 1137 LAMININ G-Like 3.  
FT DOMAIN 1139 1175 EGF-Like 11.  
FT DOMAIN 1177 1212 EGF-Like 12.  
FT DOMAIN 1214 1250 EGF-Like 13.  
FT DOMAIN 1255 1295 EGF-Like 14.  
FT DOMAIN 1297 1333 EGF-Like 15.  
FT DOMAIN 1333 1333 EGF-Like 15, CALCIUM-BINDING (POTENTIAL).  
FT DISULFID 74 85 POTENTIAL.  
FT DISULFID 79 96 POTENTIAL.  
FT DISULFID 98 107 POTENTIAL.  
FT DISULFID 114 125 POTENTIAL.  
FT DISULFID 119 134 POTENTIAL.  
FT DISULFID 136 145 POTENTIAL.  
FT DISULFID 152 163 POTENTIAL.  
FT DISULFID 157 172 POTENTIAL.  
FT DISULFID 174 183 POTENTIAL.  
FT DISULFID 190 201 POTENTIAL.  
FT DISULFID 195 210 POTENTIAL.  
FT DISULFID 212 221 POTENTIAL.  
FT DISULFID 228 239 POTENTIAL.  
FT DISULFID 233 248 POTENTIAL.  
FT DISULFID 250 259 POTENTIAL.  
FT DISULFID 305 316 POTENTIAL.  
FT DISULFID 310 325 POTENTIAL.  
FT DISULFID 327 336 POTENTIAL.  
FT DISULFID 343 354 POTENTIAL.  
FT DISULFID 348 383 POTENTIAL.  
FT DISULFID 385 394 POTENTIAL.  
FT DISULFID 445 456 POTENTIAL.  
FT DISULFID 450 469 POTENTIAL.  
FT DISULFID 471 480 POTENTIAL.  
FT DISULFID 476 687 POTENTIAL.  
FT DISULFID 681 696 POTENTIAL.  
FT DISULFID 698 707 POTENTIAL.

FT DISULFID 891 902 POTENTIAL.  
FT DISULFID 896 911 POTENTIAL.  
FT DISULFID 913 922 POTENTIAL.  
FT DISULFID 1143 1154 POTENTIAL.  
FT DISULFID 1148 1163 POTENTIAL.  
FT DISULFID 1165 1174 POTENTIAL.  
FT DISULFID 1181 1191 POTENTIAL.  
FT DISULFID 1186 1200 POTENTIAL.  
FT DISULFID 1202 1211 POTENTIAL.  
FT DISULFID 1218 1229 POTENTIAL.  
FT DISULFID 1223 1238 POTENTIAL.  
FT DISULFID 1240 1249 POTENTIAL.  
FT DISULFID 1259 1274 POTENTIAL.  
FT DISULFID 1268 1283 POTENTIAL.  
FT DISULFID 1285 1294 POTENTIAL.  
FT DISULFID 1301 1312 POTENTIAL.  
FT DISULFID 1306 1321 POTENTIAL.  
FT DISULFID 1323 1332 POTENTIAL.  
FT CARBOHYD 30 30 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 215 215 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 287 287 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 313 313 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 322 322 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 418 418 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 427 427 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 433 453 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 657 657 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 757 757 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 871 871 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 880 880 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 968 968 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 975 975 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1000 1000 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1190 1190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1243 1243 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1265 1265 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1273 1273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 161 161 A -> V (IN RP12).  
FT VARIANT 250 250 C -> W (IN RP12).  
FT VARIANT 745 745 /FTIG-VAR\_011642.  
FT VARIANT 764 764 T -> M (IN RP12).  
FT VARIANT 764 764 /FTIG-VAR\_011643.  
FT VARIANT 948 948 R -> C (IN RP12).  
FT VARIANT 948 948 C -> Y (IN RP12).  
FT VARIANT 1041 1041 /FTIG-VAR\_011645.  
FT VARIANT 1041 1041 M -> T (IN RP12).  
FT VARIANT 1071 1071 /FTIG-VAR\_011646.  
FT VARIANT 1071 1071 L -> P (IN RP12).  
FT VARIANT 1100 1100 /FTIG-VAR\_011647.  
FT VARIANT 1100 1100 I -> R (IN LCA).  
FT VARIANT 1181 1181 /FTIG-VAR\_011648.  
FT VARIANT 1181 1181 C -> R (IN RP WITH COAT-LIKE EXUDATIVE VASCULOPATHY).  
FT SEQUENCE 1376 AA; 151412 MM; F380DEF2AA046A2FE CRC64;  
FT /FTIG-VAR\_011649.  
FT /FTIG-VAR\_011649.

Query Match 38.6%; Score 83; DB 1; Length 1376;  
Best Local Similarity 39.4%; Pred. No. 0.0081;  
Matches 13; Conservative 8; Mismatches 10; Indels 2; Gaps 1;

QY 4 WS--CSGHECEVETIINHTCNCQDVGYYGQOOL 34

DB 306 WSKPCHNNATCEDSVNNTYCHCWPGYTGAQCEI 338

RESULT 36

CADN\_DROME  
 ID CADN\_DROME STANDARD: PRT: 3097 AA.  
 AC 015943; Q9VJB7;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Neural-cadherin precursor (Cadherin-N protein) (DN-cadherin).  
 GN CAN OR CG7100.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 CX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Head, and Embryo;  
 RX MEDLINE=97388431; PubMed=9247265;  
 RA Iwai Y., Usui T., Hirano S., Steward R., Takeichi M., Uemura T.;  
 RT "Axon patterning requires DN-cadherin, a novel neuronal adhesion  
 receptor, in the Drosophila embryonic CNS.";  
 RL Neuron 19:77-89(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Ciesley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Paulis B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foaier C., Gabrielian A.E., Garcia N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ileguam C.,  
 RA Jalali M., Kalush F., Kapen G.H., Ke Z., Kemison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venier E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP INTERACTION WITH ARM.  
 RX MEDLINE=98298928; PubMed=9635189;  
 RA Loureiro J., Peifer M.;  
 RT "Roles of Armadillo, a Drosophila catenin, during central nervous  
 system development.";  
 RL Curr. Biol. 8:622-632(1998).  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE

CC SORTING OF HETEROGENEOUS CELL TYPES. MAY ASSOCIATE WITH ARM NEURAL  
 CC ISOFORM AND PARTICIPATE IN THE TRANSMISSION OF DEVELOPMENTAL  
 CC INFORMATION.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (potential).  
 CC -1- TISSUE SPECIFICITY: IN THE EMBRYO, THE PROTEIN FIRST APPEARS IN  
 CC THE MESODERM AT STAGE 9 AND IS PRESENT IN THE PROBLASTS AND MUSCLE  
 CC FIBERS BY STAGE 12 AND STAGE 14, RESPECTIVELY. AT STAGE 12 THE  
 CC PROTEIN IS ALSO LOCATED IN THE AXONS OF THE ENTIRE CNS, BUT NOT IN  
 CC THE GLIAL CELLS. IN THIRD INSTAR LARVAE PROTEIN IS EXPRESSED IN  
 CC THE CNS NEUROPILE, PHOTORECEPTOR AXONS AND PRECURSORS OF ADULT  
 CC MUSCLES.  
 CC -1- SIMILARITY: CONTAINS 16 CADHERIN DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 3 EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 2 LAMININ G-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB002397; BAA22151.1; -;  
 CC EMBL: AE003656; AAF53635.1; -;  
 CC HSSP: P00740; IIXA.  
 CC FLYBASE: FBgn0015609; Cadn.  
 CC InterPro: IPR002126; Cadherin.  
 CC InterPro: IPR000233; Cadherin\_C-term.  
 CC InterPro: IPR000561; EGF-like.  
 CC InterPro: IPR000742; EGF\_2.  
 CC InterPro: IPR001881; EGF\_Ca.  
 CC InterPro: IPR001791; Laminin\_G.  
 CC Pfam: PF00028; cadherin; 14.  
 CC Pfam: PF01049; Cadherin\_C-term; 1.  
 CC Pfam: PF00008; EGF; 3.  
 CC Pfam: PF00054; Laminin\_G; 2.  
 CC PRINTS: PR00205; CADHERIN.  
 CC SMART: SM00112; CA; 16.  
 CC SMART: SM00179; EGF\_CA; 1.  
 CC SMART: SM00017; EGF-like; 3.  
 CC SMART: SM00282; LamiG; 2.  
 CC PROSITE: PS00232; CADHERIN\_1; 9.  
 CC PROSITE: PS0268; CADHERIN\_2; 16.  
 CC PROSITE: PS00022; EGF\_1; 3.  
 CC PROSITE: PS01186; EGF\_2; 3.  
 CC PROSITE: PS50025; LAM\_G\_DOMAIN; 2.  
 CC Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 CC Signal; EGF-like domain.  
 CC SIGNAL 1 36  
 CC PROPEP 37 ?  
 CC CHAIN ? 1454  
 CC TRANSMEM 1475  
 CC DOMAIN 1476 3097  
 CC DOMAIN 181 305  
 CC DOMAIN 430 543  
 CC DOMAIN 554 651  
 CC DOMAIN 660 756  
 CC DOMAIN 766 858  
 CC DOMAIN 867 968  
 CC DOMAIN 978 1078  
 CC DOMAIN 1087 1183  
 CC DOMAIN 1193 1299  
 CC DOMAIN 1307 1414  
 CC DOMAIN 1423 1514  
 CC DOMAIN 1523 1630  
 CC DOMAIN 1639 1742  
 CC DOMAIN 1749 1861  
 CC DOMAIN 1870 1966  
 CC DOMAIN 1974 2085  
 CC DOMAIN 2346 2377  
 CC DOMAIN 2379 2585  
 CC NEURAL-CADHERIN.  
 CC EXTRACELLULAR (POTENTIAL).  
 CC POTENTIAL.  
 CC CYTOPLASMIC (POTENTIAL).  
 CC CADHERIN 1.  
 CC CADHERIN 2.  
 CC CADHERIN 3.  
 CC CADHERIN 4.  
 CC CADHERIN 5.  
 CC CADHERIN 6.  
 CC CADHERIN 7.  
 CC CADHERIN 8.  
 CC CADHERIN 9.  
 CC CADHERIN 10.  
 CC CADHERIN 11.  
 CC CADHERIN 12.  
 CC CADHERIN 13.  
 CC CADHERIN 14.  
 CC CADHERIN 15.  
 CC CADHERIN 16.  
 CC EGF-LIKE 1.  
 CC LAMININ G-LIKE 1.



```

FT DOMAIN 1346 1382 EGF-LIKE 35.
FT DOMAIN 1385 1423 EGF-LIKE 36.
FT REPEAT 1486 1486 LIN/NOTCH 1.
FT REPEAT 1487 1520 LIN/NOTCH 2.
FT REPEAT 1521 1561 LIN/NOTCH 3.
FT REPEAT 1667 1910 ANK 1.
FT REPEAT 1915 1944 ANK 2.
FT REPEAT 1948 1978 ANK 3.
FT REPEAT 1982 2011 ANK 4.
FT REPEAT 2015 2044 ANK 5.
FT REPEAT 2048 2077 ANK 6.
FT DOMAIN 2265 2276 POLY-GLN (OPA-REPEAT).
FT DISULFID 25 35 BY SIMILARITY.
FT DISULFID 29 45 BY SIMILARITY.
FT DISULFID 47 56 BY SIMILARITY.
FT DISULFID 62 73 BY SIMILARITY.
FT DISULFID 67 86 BY SIMILARITY.
FT DISULFID 88 97 BY SIMILARITY.
FT DISULFID 105 116 BY SIMILARITY.
FT DISULFID 110 126 BY SIMILARITY.
FT DISULFID 128 137 BY SIMILARITY.
FT DISULFID 143 154 BY SIMILARITY.
FT DISULFID 148 163 BY SIMILARITY.
FT DISULFID 165 174 BY SIMILARITY.
FT DISULFID 181 194 BY SIMILARITY.
FT DISULFID 188 203 BY SIMILARITY.
FT DISULFID 205 214 BY SIMILARITY.
FT DISULFID 221 232 BY SIMILARITY.
FT DISULFID 226 242 BY SIMILARITY.
FT DISULFID 244 253 BY SIMILARITY.
FT DISULFID 260 271 BY SIMILARITY.
FT DISULFID 265 280 BY SIMILARITY.
FT DISULFID 282 291 BY SIMILARITY.
FT DISULFID 298 311 BY SIMILARITY.
FT DISULFID 305 320 BY SIMILARITY.
FT DISULFID 322 331 BY SIMILARITY.
FT DISULFID 338 349 BY SIMILARITY.
FT DISULFID 343 358 BY SIMILARITY.
FT DISULFID 360 369 BY SIMILARITY.
FT DISULFID 375 386 BY SIMILARITY.
FT DISULFID 380 397 BY SIMILARITY.
FT DISULFID 399 408 BY SIMILARITY.
FT DISULFID 415 428 BY SIMILARITY.
FT DISULFID 422 437 BY SIMILARITY.
FT DISULFID 439 448 BY SIMILARITY.
FT DISULFID 455 466 BY SIMILARITY.
FT DISULFID 460 475 BY SIMILARITY.
FT DISULFID 477 486 BY SIMILARITY.
FT DISULFID 493 503 BY SIMILARITY.
FT DISULFID 498 512 BY SIMILARITY.
FT DISULFID 514 523 BY SIMILARITY.
FT DISULFID 530 541 BY SIMILARITY.
FT DISULFID 535 550 BY SIMILARITY.
FT DISULFID 552 561 BY SIMILARITY.
FT DISULFID 568 578 BY SIMILARITY.
FT DISULFID 573 587 BY SIMILARITY.
FT DISULFID 589 598 BY SIMILARITY.
FT DISULFID 605 616 BY SIMILARITY.
FT DISULFID 610 625 BY SIMILARITY.
FT DISULFID 627 636 BY SIMILARITY.
FT DISULFID 643 653 BY SIMILARITY.
FT DISULFID 648 662 BY SIMILARITY.
FT DISULFID 664 673 BY SIMILARITY.
FT DISULFID 680 691 BY SIMILARITY.
FT DISULFID 685 700 BY SIMILARITY.
FT DISULFID 702 711 BY SIMILARITY.
FT DISULFID 716 728 BY SIMILARITY.
FT DISULFID 723 737 BY SIMILARITY.
FT DISULFID 739 748 BY SIMILARITY.
FT DISULFID 755 766 BY SIMILARITY.
FT DISULFID 760 775 BY SIMILARITY.
FT DISULFID 777 786 BY SIMILARITY.
FT DISULFID 793 804 BY SIMILARITY.

```

```

FT DISULFID 798 813 BY SIMILARITY.
FT DISULFID 815 824 BY SIMILARITY.
FT DISULFID 831 842 BY SIMILARITY.
FT DISULFID 836 853 BY SIMILARITY.
FT DISULFID 855 864 BY SIMILARITY.
FT DISULFID 871 882 BY SIMILARITY.
FT DISULFID 876 891 BY SIMILARITY.
FT DISULFID 893 902 BY SIMILARITY.
FT DISULFID 909 920 BY SIMILARITY.
FT DISULFID 914 929 BY SIMILARITY.
FT DISULFID 931 940 BY SIMILARITY.
FT DISULFID 947 958 BY SIMILARITY.
FT DISULFID 952 967 BY SIMILARITY.
FT DISULFID 969 978 BY SIMILARITY.
FT DISULFID 1023 1034 BY SIMILARITY.
FT DISULFID 1028 1043 BY SIMILARITY.
FT DISULFID 1045 1054 BY SIMILARITY.

Query Match 38.1%; Score 82; DB 1; Length 2437;
Best Local Similarity 32.4%; Pred. No. 0.018;
Matches 11; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

Oy 1 CQPMSCSGHGECEIINNHOTCDVGYGPQCL 34
Db 947 CVSAPCRNGNCCTDCVNSYTCSCPAGSGINCEI 980

RESULT 38
ID FA12_HUMAN STANDARD; PRT; 615 AA.
AC P00748; P78339;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coagulation factor XII precursor (EC 3.4.21.38) (Hageman factor)
DE (Haf).
GN F12.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP MEDLINE=88007593; PubMed=2888762;
RX Cool D.E., McGilivray R.T.A.;
RA "Characterization of the human blood coagulation factor XII gene.
RT Intron/exon gene organization and analysis of the 5'-flanking
RT region.";
RL J. Biol. Chem. 262:13662-13673(1987).
RN [2]
RP SEQUENCE OF 4-615 FROM N.A.
RX MEDLINE=86176794; PubMed=3754331;
RA Tripodi M., Citarella F., Guida S., Galeffi P., Fantoni A.,
RA Cortese R.;
RT "cDNA sequence coding for human coagulation factor XII (Hageman).";
RT Nucleic Acids Res. 14:3146-3146(1986).
RN [3]
RP SEQUENCE OF 14-615 FROM N.A.
RX MEDLINE=86033830; PubMed=3877053;
RA Cool D.E., Edgell C.-J.S., Louie G.V., Zoller M.J., Brayer G.D.,
RA McGilivray R.T.A.;
RT "Characterization of human blood coagulation factor XII cDNA.
RT Prediction of the primary structure of factor XII and the tertiary
RT structure of beta-factor Xlla.";
RL J. Biol. Chem. 260:13666-13676(1985).
RN [4]
RP SEQUENCE OF 146-615 FROM N.A.
RX MEDLINE=86216049; PubMed=3011063;
RA Que B.G., Davie E.W.;
RT "Characterization of a cDNA coding for human factor XII (Hageman
RT factor).";
RL Biochemistry 25:1525-1528(1986).
RN [5]

```

RP SEQUENCE OF 20-379.  
RX MEDLINE-85182674; PubMed-3886654;  
RA McMullen B.A., Fujikawa K.;  
RT "Amino acid sequence of the heavy chain of human alpha-factor XIIa  
RT (activated Hageman factor).";  
RL J. Biol. Chem. 260:5328-5341(1985).  
RN [6]  
RP SEQUENCE OF 354-362 AND 373-615.  
RX MEDLINE-83291041; PubMed-6604055;  
RA Fujikawa K., McMullen B.A.;  
RT "Amino acid sequence of human beta-factor XIIa.";  
RL J. Biol. Chem. 258:10924-10933(1983).  
RN [7]  
RP TISSUE-Blood:  
RC SEQUENCE OF 561-615 FROM N.A.  
RX MEDLINE-96133302; PubMed-8528215;  
RA Schloesser M., Hofferbert S., Bartz U., Lutze G., Lammle B., Engel W.;  
RT "The novel acceptor splice site mutation 11396(G->A) in the factor  
RT XII gene causes a truncated transcript in cross-reacting material  
RT negative patients.";  
RL Hum. Mol. Genet. 4:1235-1237(1995).  
RN [8]  
RP VARIANT WASHINGTON DC.  
RX MEDLINE-90046788; PubMed-2510163;  
RA Miyata T., Kawabata S.-I., Iwanaga S., Takahashi I., Alying B.,  
RA Saito H.;  
RT "Coagulation factor XII (Hageman factor) Washington D.C.: inactive  
RT factor XIIa results from Cys-571->Ser substitution.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:8319-8322(1989).  
RN [9]  
RP VARIANT LOCARNO.  
RX MEDLINE-94325559; PubMed-8049433;  
RA Hovington J.K., Schaller J., Stricker H., Wulllemann W.A., Furian M.,  
RA Laemmle B.;  
RT "Coagulation factor XII Locarno: the functional defect is caused by  
RT the amino acid substitution Arg-353->Pro leading to loss of a  
RT kallikrein cleavage site.";  
RL Blood 84:1173-1181(1994).  
RN [10]  
RP CARBOHYDRATE-LINKAGE SITE THR-109.  
RX MEDLINE-92184750; PubMed-1544894;  
RA Harris R.J., Ling V.T., Spellman M.W.;  
RT "O-linked fucose is present in the first epidermal growth factor  
RT domain of factor XII but not protein C.";  
RL J. Biol. Chem. 267:5102-5107(1992).  
RN [11]  
RP FUNCTION: FACTOR XII IS A SERUM GLYCOPROTEIN THAT PARTICIPATES IN  
CC THE INITIATION OF BLOOD COAGULATION, FIBRINOLYSIS, AND THE  
CC GENERATION OF BRADYKININ AND ANGIOTENSIN.  
CC [12]  
CC CATALYTIC ACTIVITY: Cleaves selectively Arg-|-Ile bonds in factor  
CC VII to form factor VIIa and factor XI to form factor Xla.  
CC [13]  
CC PTM: O- AND N-GLYCOSYLATED.  
CC [14]  
CC DISEASE: DEFECTS IN F12 DO NOT CAUSE ANY CLINICAL SYMPTOMS. THE  
CC SOLE EFFECT IS THAT WHOLE-BLOOD CLOTTING TIME IS PROLONGED.  
CC [15]  
CC MISCELLANEOUS: FACTOR XII, PREKALLIKREIN, AND HMW KININOGEN FORM A  
CC COMPLEX BOUND TO AN ANIONIC SURFACE. PREKALLIKREIN IS CLEAVED BY  
CC FACTOR XII TO FORM KALLIKREIN, WHICH THEN CLEAVES FACTOR XII FIRST  
CC TO ALPHA-FACTOR XIIa AND THEN TO BETA-FACTOR XIIa. ALPHA-FACTOR  
CC XIIa ACTIVATES FACTOR XI TO FACTOR Xla.  
CC [16]  
CC SIMILARITY: CONTAINS 2 EGF-LIKE DOMAINS.  
CC [17]  
CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE I DOMAIN.  
CC [18]  
CC SIMILARITY: CONTAINS 1 FIBRONECTIN TYPE II DOMAIN.  
CC [19]  
CC SIMILARITY: CONTAINS 1 KRINGLE DOMAIN.  
CC [20]  
CC SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPsin FAMILY.  
CC [21]  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

DR EMBL: M1315; AAA70225.1; -;  
DR EMBL: M1723; AAA51986.1; -;  
DR EMBL: M17466; AAB59490.1; -;  
DR EMBL: M17464; AAB59490.1; JOINED.  
DR EMBL: M17465; AAB59490.1; JOINED.  
DR EMBL: M13147; AAA70224.1; -;  
DR EMBL: U71274; AAB51203.1; -;  
DR PIR: A29411; KFHU12.  
DR HSSP: P00750; 1RTF.  
DR MEROPS: S01.211; -;  
DR MIM: 234000; -;  
DR InterPro: IPR001314; Chymotrypsin.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000562; FN\_Type\_II.  
DR InterPro: IPR000001; kringle.  
DR InterPro: IPR001284; trypsin.  
DR InterPro: IPR000083; fibrinectin\_type\_1.  
DR Pfam: PF00008; EGF\_2.  
DR Pfam: PF00039; fn1; 1.  
DR Pfam: PF00040; fn2; 1.  
DR Pfam: PF00051; kringle; 1.  
DR Pfam: PF00089; trypsin; 1.  
DR PRINTS: PR00722; CHYMOTRYPSIN.  
DR PRINTS: PR00013; ENTPEIT.  
DR PRINTS: PR00018; KRINGLE.  
DR ProDom: PD00995; FN\_Type\_II; 1.  
DR SMART: SM00181; EGF; 2.  
DR SMART: SM00058; FN1; 1.  
DR SMART: SM00059; FN2; 1.  
DR SMART: SM00130; KR; 1.  
DR SMART: SM00020; Tryp\_Spc; 1.  
DR PROSITE: PS00022; EGF\_1; 2.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR PROSITE: PS01253; FIBRONECTIN\_1; 1.  
DR PROSITE: PS00023; FIBRONECTIN\_2; 1.  
DR PROSITE: PS00021; KRINGLE\_1; 1.  
DR PROSITE: PS00070; KRINGLE\_2; 1.  
DR PROSITE: PS00240; TRYPSIN\_DOM; 1.  
DR PROSITE: PS00134; TRYPSIN\_HIS; 1.  
DR PROSITE: PS00135; TRYPSIN\_SER; 1.  
DR KMW: Hydrolyse; Fibrinolysis; Signal; EGF-like domain; Repeat; zymogen;  
KW Disease mutation.  
KW SIGNAL.  
FT CHAIN 1 19  
FT CHAIN 20 372  
FT CHAIN 373 615  
FT CHAIN 354 362  
FT CHAIN 373 615  
FT CHAIN 47 88  
FT DOMAIN 94 131  
FT DOMAIN 133 173  
FT DOMAIN 174 210  
FT DOMAIN 217 295  
FT DOMAIN 296 349  
FT DOMAIN 373 615  
FT CARBOHYD 109  
FT CARBOHYD 249  
FT CARBOHYD 259 299  
FT CARBOHYD 305 305  
FT CARBOHYD 308 308  
FT CARBOHYD 328 328  
FT CARBOHYD 329 329  
FT CARBOHYD 337 337  
FT CARBOHYD 412 412  
FT ACT\_SITE 461 461  
FT ACT\_SITE 563 563  
FT ACT\_SITE 563 563  
FT DISULFID 98 110  
FT DISULFID 104 119  
FT DISULFID 121 130  
FT DISULFID 135 163  
FT DISULFID 161 170  
FT DISULFID 178 189  
FT DISULFID 183 198  
FT DISULFID 183 198  
ALPHA-FACTOR XIIa HEAVY CHAIN.  
BETA-FACTOR XIIa LIGHT CHAIN.  
BETA-FACTOR XIIa PART 1.  
BETA-FACTOR XIIa PART 2.  
FIBRONECTIN TYPE-I.  
EGF-LIKE 1.  
EGF-LIKE 2.  
FIBRONECTIN TYPE-1.  
KRINGLE.  
PRO-RICH.  
SERINE PROTEASE.  
O-LINKED (PUC).  
N-LINKED (GLCNAC. . .).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
O-LINKED (POTENTIAL).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
CHARGE RELAY SYSTEM (BY SIMILARITY).  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.  
BY SIMILARITY.



ID	NTCI_HUMAN	STANDARD:	PRT:	2444 AA.
AC	P46531;			
DT	01-NOV-1995 (rel. 32, Created)			
DT	01-NOV-1995 (rel. 32, Last sequence update)			
DT	16-OCT-2001 (rel. 40, Last annotation update)			
DE	Neurogenic locus notch protein homolog 1 precursor (Translocation-associated notch protein TAN-1) (Fragment).			
GN	NOTCH1 OR TAN1.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91347367; PubMed=1831692;			
RA	Ellisen L.W., Bird J., West D.C., Soreng A.L., Reynolds T.C.,			
RA	Smith S.D., Sklar J.;			
RT	"TAN-1, the human homolog of the Drosophila notch gene, is broken by			
RT	chromosomal translocations in T lymphoblastic neoplasms.";			
RL	Cell 66:649-661(1991).			
CC	-1- FUNCTION: MAY BE IMPORTANT FOR NORMAL LYMPHOCYTE FUNCTION. IN			
CC	ALTERED FORM, MAY CONTRIBUTE TO TRANSFORMATION OR PROGRESSION			
CC	IN SOME T-CELL NEOPLASMS.			
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.			
CC	-1- TISSUE SPECIFICITY: IN FETAL TISSUES MOST ABUNDANT IN SPLEEN,			
CC	BRAIN STEM AND LUNG. ALSO PRESENT IN MOST ADULT TISSUES WHERE IT			
CC	IS FOUND MAINLY IN LYMPHOID TISSUES.			
CC	-1- SIMILARITY: HIGH, WITH OTHER NOTCH-TYPE PROTEINS.			
CC	-1- SIMILARITY: CONTAINS 36 EGF-LIKE DOMAINS.			
CC	-1- SIMILARITY: CONTAINS 5 ANK REPEATS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).			
CC	-----			
DR	EMBL; M73980; AAA0614.1; .			
DR	HSSP; P00740; IIXA.			
DR	MIM; 190198; .			
DR	Interpro: IPR002110; ANK.			
DR	Interpro: IPR000152; Asx_hydroxyl.			
DR	Interpro: IPR000561; EGF-like.			
DR	Interpro: IPR000742; EGF-2.			
DR	Interpro: IPR001881; EGF_Ca.			
DR	Interpro: IPR000800; Notch.			
DR	Pfam; PF00023; ank; 6.			
DR	Pfam; PF00008; EGF; 36.			
DR	Pfam; PF00066; notch; 3.			
DR	SMART; SM00248; ANK; 5.			
DR	SMART; SM00179; EGF_CA; 22.			
DR	SMART; SM00001; EGF-like; 13.			
DR	SMART; SM00004; NL; 2.			
DR	PROSITE; PS50086; ANK_REPEAT; 4.			
DR	PROSITE; PS50297; ANK_REPEAT_REGION; 1.			
DR	PROSITE; PS00010; ASX_HYDROXYL; 20.			
DR	PROSITE; PS00022; EGF_1; 34.			
DR	PROSITE; PS01186; EGF_2; 26.			
DR	PROSITE; PS01187; EGF_CA; 18.			
KW	Differentiation; Neurogenesis; Repeat; ANK repeat; EGF-like domain;			
KW	Transmembrane; Signal; Glycoprotein.			
FT	SIGNAL	1	18	POTENTIAL.
FT	CHAIN	19	>2444	NEUROGENIC LOCUS NOTCH PROTEIN HOMOLOG 1.
FT	DOMAIN	19	1736	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	1737	1757	POTENTIAL.
FT	DOMAIN	1758	>2444	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	20	58	EGF-LIKE 1.
FT	DOMAIN	59	99	EGF-LIKE 2.
FT	DOMAIN	102	139	EGF-LIKE 3.
FT	DOMAIN	140	176	EGF-LIKE 4.
FT	DOMAIN	178	216	EGF-LIKE 5.
FT	DOMAIN	218	255	EGF-LIKE 6.
FT	DOMAIN	257	293	EGF-LIKE 7.
FT	DOMAIN	295	333	EGF-LIKE 8.
FT	DOMAIN	335	371	EGF-LIKE 9.
FT	DOMAIN	372	410	EGF-LIKE 10.
FT	DOMAIN	412	450	EGF-LIKE 11.
FT	DOMAIN	452	488	EGF-LIKE 12.
FT	DOMAIN	490	526	EGF-LIKE 13.
FT	DOMAIN	528	564	EGF-LIKE 14.
FT	DOMAIN	566	601	EGF-LIKE 15.
FT	DOMAIN	603	639	EGF-LIKE 16.
FT	DOMAIN	641	676	EGF-LIKE 17.
FT	DOMAIN	678	714	EGF-LIKE 18.
FT	DOMAIN	716	751	EGF-LIKE 19.
FT	DOMAIN	753	789	EGF-LIKE 20.
FT	DOMAIN	791	827	EGF-LIKE 21.
FT	DOMAIN	829	868	EGF-LIKE 22.
FT	DOMAIN	870	906	EGF-LIKE 23.
FT	DOMAIN	908	944	EGF-LIKE 24.
FT	DOMAIN	946	982	EGF-LIKE 25.
FT	DOMAIN	984	1020	EGF-LIKE 26.
FT	DOMAIN	1022	1058	EGF-LIKE 27.
FT	DOMAIN	1060	1096	EGF-LIKE 28.
FT	DOMAIN	1098	1144	EGF-LIKE 29.
FT	DOMAIN	1146	1182	EGF-LIKE 30.
FT	DOMAIN	1184	1220	EGF-LIKE 31.
FT	DOMAIN	1222	1266	EGF-LIKE 32.
FT	DOMAIN	1268	1306	EGF-LIKE 33.
FT	DOMAIN	1308	1347	EGF-LIKE 34.
FT	DOMAIN	1349	1385	EGF-LIKE 35.
FT	DOMAIN	1388	1427	EGF-LIKE 36.
FT	REPEAT	1446	1481	LIN/NOTCH 1.
FT	REPEAT	1482	1523	LIN/NOTCH 2.
FT	REPEAT	1524	1563	LIN/NOTCH 3.
FT	REPEAT	1598	1637	ANK 1.
FT	REPEAT	1638	1677	ANK 2.
FT	REPEAT	1678	1717	ANK 3.
FT	REPEAT	1718	1757	ANK 4.
FT	REPEAT	1758	1797	ANK 5.
FT	DOMAIN	1798	1837	POLY-VAL.
FT	DOMAIN	1838	1877	POLY-ARG.
FT	DOMAIN	1878	1917	POLY-PRO.
FT	DOMAIN	1918	1957	POLY-ALA.
FT	DOMAIN	1958	1997	POLY-GLU.
FT	DOMAIN	1998	2037	POLY-GLY.
FT	DOMAIN	2038	2077	POLY-GLN.
FT	DOMAIN	2078	2117	POLY-PRO.
FT	DISULFID	2118	2157	BY SIMILARITY.
FT	DISULFID	2158	2197	BY SIMILARITY.
FT	DISULFID	2198	2237	BY SIMILARITY.
FT	DISULFID	2238	2277	BY SIMILARITY.
FT	DISULFID	2278	2317	BY SIMILARITY.
FT	DISULFID	2318	2357	BY SIMILARITY.
FT	DISULFID	2358	2397	BY SIMILARITY.
FT	DISULFID	2398	2437	BY SIMILARITY.
FT	DISULFID	2438	2477	BY SIMILARITY.
FT	DISULFID	2478	2517	BY SIMILARITY.
FT	DISULFID	2518	2557	BY SIMILARITY.
FT	DISULFID	2558	2597	BY SIMILARITY.
FT	DISULFID	2598	2637	BY SIMILARITY.
FT	DISULFID	2638	2677	BY SIMILARITY.
FT	DISULFID	2678	2717	BY SIMILARITY.
FT	DISULFID	2718	2757	BY SIMILARITY.
FT	DISULFID	2758	2797	BY SIMILARITY.
FT	DISULFID	2798	2837	BY SIMILARITY.
FT	DISULFID	2838	2877	BY SIMILARITY.
FT	DISULFID	2878	2917	BY SIMILARITY.
FT	DISULFID	2918	2957	BY SIMILARITY.
FT	DISULFID	2958	2997	BY SIMILARITY.
FT	DISULFID	2998	3037	BY SIMILARITY.
FT	DISULFID	3038	3077	BY SIMILARITY.
FT	DISULFID	3078	3117	BY SIMILARITY.
FT	DISULFID	3118	3157	BY SIMILARITY.
FT	DISULFID	3158	3197	BY SIMILARITY.
FT	DISULFID	3198	3237	BY SIMILARITY.
FT	DISULFID	3238	3277	BY SIMILARITY.
FT	DISULFID	3278	3317	BY SIMILARITY.
FT	DISULFID	3318	3357	BY SIMILARITY.
FT	DISULFID	3358	3397	BY SIMILARITY.

FT	DISULEID	344	359	BY SIMILARITY.
FT	DISULEID	361	370	BY SIMILARITY.
FT	DISULEID	376	387	BY SIMILARITY.
FT	DISULEID	381	398	BY SIMILARITY.
FT	DISULEID	400	409	BY SIMILARITY.
FT	DISULEID	416	429	BY SIMILARITY.
FT	DISULEID	423	438	BY SIMILARITY.
FT	DISULEID	440	449	BY SIMILARITY.
FT	DISULEID	456	467	BY SIMILARITY.
FT	DISULEID	461	476	BY SIMILARITY.
FT	DISULEID	478	487	BY SIMILARITY.
FT	DISULEID	494	505	BY SIMILARITY.
FT	DISULEID	499	514	BY SIMILARITY.
FT	DISULEID	516	525	BY SIMILARITY.
FT	DISULEID	532	543	BY SIMILARITY.
FT	DISULEID	537	552	BY SIMILARITY.
FT	DISULEID	554	563	BY SIMILARITY.
FT	DISULEID	570	580	BY SIMILARITY.
FT	DISULEID	575	589	BY SIMILARITY.
FT	DISULEID	591	600	BY SIMILARITY.
FT	DISULEID	607	618	BY SIMILARITY.
FT	DISULEID	612	627	BY SIMILARITY.
FT	DISULEID	629	638	BY SIMILARITY.
FT	DISULEID	645	655	BY SIMILARITY.
FT	DISULEID	650	664	BY SIMILARITY.
FT	DISULEID	666	675	BY SIMILARITY.
FT	DISULEID	682	693	BY SIMILARITY.
FT	DISULEID	687	702	BY SIMILARITY.
FT	DISULEID	704	713	BY SIMILARITY.
FT	DISULEID	720	730	BY SIMILARITY.
FT	DISULEID	725	739	BY SIMILARITY.
FT	DISULEID	741	750	BY SIMILARITY.
FT	DISULEID	757	768	BY SIMILARITY.
FT	DISULEID	762	777	BY SIMILARITY.
FT	DISULEID	779	788	BY SIMILARITY.
FT	DISULEID	795	806	BY SIMILARITY.
FT	DISULEID	800	815	BY SIMILARITY.
FT	DISULEID	817	826	BY SIMILARITY.
FT	DISULEID	833	844	BY SIMILARITY.
FT	DISULEID	838	855	BY SIMILARITY.
FT	DISULEID	857	867	BY SIMILARITY.
FT	DISULEID	874	885	BY SIMILARITY.
FT	DISULEID	879	894	BY SIMILARITY.
FT	DISULEID	896	905	BY SIMILARITY.
FT	DISULEID	912	923	BY SIMILARITY.
FT	DISULEID	917	932	BY SIMILARITY.
FT	DISULEID	934	943	BY SIMILARITY.
FT	DISULEID	988	999	BY SIMILARITY.
FT	DISULEID	993	1008	BY SIMILARITY.
FT	DISULEID	1010	1019	BY SIMILARITY.
FT	DISULEID	1026	1037	BY SIMILARITY.
FT	DISULEID	1031	1046	BY SIMILARITY.
FT	DISULEID	1048	1057	BY SIMILARITY.
FT	DISULEID	1064	1075	BY SIMILARITY.

Query Match 37.7%; Score 81; DB 1; Length 2444;  
Best Local Similarity 45.5%; Pred. No. 0.024;

Matches 15; Conservative 3; Mismatches 15; Indels 0; Gaps 0;

OY 1 CQPMSCGHECVETINNHTCNCVGYGPOCQ 33  
Db 988 CTSSCFMGTCVDGINSFTCLCPGFGSYCQ 1020

Search completed: September 7, 2002, 10:23:43  
Job time: 536 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:15:25 : Search time 36.49 Seconds  
(without alignments)  
89.532 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193

Perfect score: 215

Sequence: 1 COPWSCSGHGECVEIINNHTCNCDCVGYGPOCOL 34

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

PIR-71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	205	95.3	323	1	S09702 L-selectin precurs
2	205	95.3	372	2	JC5377 L-selectin precurs
3	205	95.3	385	1	A34015 L-selectin precurs
4	191	88.8	370	2	S22124 L-selectin precurs
5	181	84.2	376	2	JC4892 L-selectin precurs
6	169	78.6	372	1	A32375 L-selectin precurs
7	159	74.0	372	2	S23936 L-selectin precurs
8	130	60.5	610	2	A35046 L-selectin precurs
9	127	59.1	646	2	JN0473 L-selectin precurs
10	126	58.6	612	2	B42755 E-selectin precurs
11	125	58.1	482	2	JC5092 E-selectin precurs
12	125	58.1	830	2	A30359 E-selectin - pig
13	124	57.7	768	2	A42755 P-selectin precurs
14	118	54.9	485	2	S36772 E-selectin - bov
15	116	54.0	551	2	I46709 endothelial leuko
16	109	50.7	768	2	I53821 P-selectin - rat
17	95	44.2	603	2	S28941 coagulation factor
18	95	44.2	710	2	T21339 hypothetical prote
19	93	43.3	2524	2	A35844 Notch protein - Af
20	90	41.9	2531	2	T31070 notch homolog - se
21	89	41.9	3051	2	S42373 hypothetical prote
22	89	41.4	1203	2	A49175 Notch B protein -
23	89	41.4	2421	2	S78549 notch3 protein - h
24	89	41.4	2471	2	A49128 cell-fate determin
25	87	40.5	461	1	JY0210 protein C (activat
26	87	40.5	728	2	I50719 C-beta-1 - chick
27	87	40.5	1064	2	A40136 fibropellin Ia - s
28	86	40.0	2139	2	A35672 crumbs protein - f
29	86	40.0	2703	1	A24420 notch protein - fr

30	85.5	39.8	473	2	A56175 adhesive plaque pr
31	85.5	39.8	3461	2	S58870 reelin precursor -
32	85	39.5	1722	2	E89753 protein FliC7.4 (1
33	85	39.5	2352	2	T30201 Notch homolog prot
34	84	39.1	372	2	T29359 hypothetical prote
35	84	39.1	2019	1	JQ1322 tenascin precursor
36	84	39.1	2201	2	A32160 tenascin-C - human
37	84	39.1	2318	2	S45306 notch 3 protein -
38	84	39.1	3623	2	T09456 intrinsic factor-B
39	83.5	38.8	1069	2	T42681 hypothetical prote
40	83.5	38.8	1523	2	T13953 MEK5 protein - ra
41	83.5	38.8	2809	2	T30213 G-cadherin - sea u
42	83	38.6	293	2	B26337 neurogenic repetit
43	83	38.6	461	1	S18994 protein C (activat
44	83	38.6	570	2	A48836 fibropellin C prec
45	83	38.6	3097	2	T00021 DN-cadherin - Fru1

#### ALIGNMENTS

RESULT 1  
S09702  
L-selectin precursor, short splice form - human  
N/Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc  
ral lymph node homing receptor Leu-8  
C/Species: Homo sapiens (man)  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 26-May-2000  
C/Accession: S09702  
R/Camerini, D.; James, S.P.; Stamenkovic, I.; Seed, B.  
Nature 342, 78-82, 1989  
A/Title: Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.  
A/Reference number: S06796; MUID:90044046  
A/Accession: S09702  
A/Status: not compared with conceptual translation  
A/Molecule type: mRNA  
A/Residues: 1-323 <CAM>  
A/Cross-references: EMBL:X17519; NID:g34344  
A/Note: this translation is not annotated in GenBank entry HSLEU8, release 111.0  
C/Comment: For an alternative splice form, see PIR:A34015.  
C/Genetics:  
A/Gene: GDB:SEL; GDB:LNHR; LSEL; LAM1; LYAM1; LAM-1  
A/Cross-references: GDB:120157; GDB:118834; OMIM:153240  
A/Map position: 1q22-1q23  
A/Introns: 14/3; 42/1; 171/1; 207/1; 269/1; 303/2  
C/Function:  
A/Description: binds with low affinity to oligosaccharides like heparan sulfate and si  
ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr  
C/Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology  
C/Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammat  
F:1-51/Domain: signal sequence #status predicted <SIG>  
F:42-168/Domain: C-type lectin homology <LCH>  
F:52-323/Product: L-selectin #status predicted <MAT>  
F:52-300/Domain: extracellular #status predicted <EXT>  
F:173-204/Domain: EGF homology <EGF>  
F:210-267/Domain: complement factor H repeat homology <FH1>  
F:272-322/Domain: complement factor H repeat homology #status atypical <FH2>  
F:301-318/Domain: transmembrane #status predicted <TM>  
F:319-322/Domain: intracellular #status predicted <INT>  
F:73-117,190,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:322/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 95.3%; Score 205; DB 1; Length 323;  
Best Local Similarity 97.0%; Pred. No. 7e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECVEIINNHTCNCDCVGYGPOCO 33  
DB 173 COPWSCSGHGECVEIINNHTCNCDCVGYGPOCO 205

RESULT 2

JC5377  
 L-selectin precursor - hamadryas baboon  
 C:Species: Papio hamadryas (hamadryas baboon)  
 C>Date: 02-Jun-1997 #sequence\_revision 12-Sep-1997 #text\_change 21-Jan-2000  
 C:Accession: JC5377; PC4315  
 R:tsurushita, N.; Fu, H.; Berg, E.L.  
 Gene 181, 219-220, 1996  
 A:Title: PCR cloning of the cDNA encoding baboon L-selectin.  
 A:Reference number: JC5377; MUID:97128794  
 A:Accession: JC5377  
 A:Molecule type: mRNA  
 A:Residues: 1-372 <TSU1>  
 A:Cross-references: GB:U52074; NID:g1326148; PIDN:AAB40903.1; PID:g1326149  
 A:Accession: PC4315  
 A:Molecule type: protein  
 A:Residues: 37-43:142-148 <TSU2>  
 C:Comment: This receptor is involved in the initial adhesive interaction between lymphoc sites of inflammation.  
 C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology;  
 F:1-38/Domain: signal sequence #status predicted <SIG>  
 F:29-155/Domain: C-type lectin homology <LCH>  
 F:39-372/Product: L-selectin #status predicted <MAT>  
 F:39-157/Domain: calcium-binding #status predicted <CAB>  
 F:160-191/Domain: EGF homology <EGF>  
 F:197-254/Domain: complement factor H repeat homology <FH1>  
 F:259-316/Domain: complement factor H repeat homology <FH2>  
 F:333-355/Domain: transmembrane #status predicted <TM>  
 F:356-372/Domain: intracellular #status predicted <INT>

Query Match 95.3%; Score 205; DB 2; Length 372;  
 Best Local Similarity 97.0%; Pred. No. 7.9e-16;  
 Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCGHGECVEIINNHTCNCVGYGPQCQ 33  
 Db 160 CQPMSCGHGECVEIINNHTCNCVGYGPQCQ 192

## RESULT 3

A34015  
 L-selectin precursor, long splice form - human  
 N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukocyte  
 ral lymph node homing receptor Leu-8  
 C:Species: Homo sapiens (man)  
 C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
 C:Accession: I55333; S06798; JI0104; A34015; A33912  
 R:Ord, D.C.; Ernst, T.J.; Zhou, L.J.; Rambaldi, A.; Spertini, O.; Griffin, J.; Tedder, T.  
 J. Biol. Chem. 265, 7760-7767, 1990  
 A:Title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (TQ1, Le  
 A:Reference number: I55333; MUID:90243637  
 A:Accession: I55333  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 14-385 <ORD>  
 A:Cross-references: GB:I32414; NID:g187259; PIDN:AAB60700.1; PID:g386860  
 R:Cammerini, D.; James, S.P.; Stamenkovic, I.; Seed, B.  
 Nature 342, 78-82, 1989  
 A:Title: Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing receptor.  
 A:Reference number: S06798; MUID:90044046  
 A:Accession: S06798

A:Molecule type: mRNA  
 A:Residues: 1-225, 'S', 227-385 <CAM>  
 A:Cross-references: EMBL:X17519; NID:g34344; PIDN:CAB43536.1; PID:g4902829  
 A:Note: this translation is not annotated in GenBank entry HSLER8, release 111.0  
 R:tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Disteche, C.M.  
 J. Exp. Med. 170, 123-133, 1989  
 A:Title: Isolation and chromosomal localization of cDNAs encoding a novel human lymphocy  
 oteins.  
 A:Reference number: JI0104; MUID:89310350  
 A:Accession: JI0104  
 A:Molecule type: mRNA  
 A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TED>

A:Cross-references: GB:X16150; NID:g34428; PIDN:CAA34275.1; PID:g34429  
 A:Note: the translated sequence in GenBank entry HSLVWML, release 111.0, differs from  
 R:Bowen, B.R.; Nguyen, T.; Lasky, L.A.  
 J. Cell Biol. 109, 421-427, 1989  
 A:Title: Characterization of a human homologue of the murine peripheral lymph node ho  
 A:Reference number: A34015; MUID:89308881  
 A:Accession: A34015  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 14-49, 'V', 51-190, 'H', 192-205, 'L', 207-226, 'F', 228-385 <ROW>  
 A:Cross-references: GB:X16070; NID:g38092; PIDN:CAA34203.1; PID:g38093  
 R:Siegelman, M.H.; Weissman, I.L.  
 Proc. Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989  
 A:Title: Human homologue of mouse lymph node homing receptor: evolutionary conservati  
 A:Reference number: A33912; MUID:89315837  
 A:Accession: A33912  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 14-205, 'L', 207-385 <SIE>  
 A:Cross-references: GB:M25280; NID:g187182; PIDN:AAC63053.1; PID:g307134  
 C:Comment: For an alternative splice form, see PIR:S09702.  
 C:Genetics:  
 A:Gene: GDB:SELL; GDB:LNHR; ISFL; LAM1; LYAM1; LAM-1  
 A:Cross-references: GDB:120157; GDB:118834; OMIM:153240  
 A:Map position: 1922-1923  
 A:Insertions: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2  
 C:Function:  
 A:Description: binds with low affinity to oligosaccharides like heparan sulfate and si  
 ment of leukocytes to areas of inflammation, and with CD162 mediates neutrophil-neutr  
 C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homolo  
 C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammat  
 F:1-51/Domain: signal sequence #status predicted <SIG>  
 F:42-168/Domain: C-type lectin homology <LCH>  
 F:52-385/Product: L-selectin #status predicted <MAT>  
 F:52-343/Domain: extracellular #status predicted <EXT>  
 F:173-204/Domain: EGF homology <EGF>  
 F:210-267/Domain: complement factor H repeat homology <FH1>  
 F:272-329/Domain: complement factor H repeat homology <FH2>  
 F:344-368/Domain: transmembrane #status predicted <TM>  
 F:369-385/Domain: intracellular #status predicted <INT>  
 F:73, 117, 190, 245, 259/Binding site: carbohydrate (asn) (covalent) #status predicted  
 F:377,380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 95.3%; Score 205; DB 1; Length 385;  
 Best Local Similarity 97.0%; Pred. No. 8.1e-16;  
 Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCGHGECVEIINNHTCNCVGYGPQCQ 33  
 Db 173 CQPMSCGHGECVEIINNHTCNCVGYGPQCQ 205

## RESULT 4

S22124  
 L-selectin precursor - bovine  
 N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
 C:Accession: S22124; A46531  
 R:Bosworth, B.T.  
 submitted to the EMBL Data Library, October 1991  
 A:Reference number: S22123  
 A:Accession: S22124  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-370 <BOS>  
 A:Cross-references: EMBL:X62882; NID:g515; PIDN:CAA44676.1; PID:g516  
 R:Witcher, B.; White, M.; Kurk, S.; Kishimoto, T.K.; Jutila, M.A.  
 Eur. J. Immunol. 22, 469-476, 1992  
 A:Title: Characterization of the bovine peripheral lymph node homing receptor: a lect  
 A:Reference number: A46531; MUID:92164727  
 A:Accession: A46531



Query Match 74.0%; Score 159; DB 2; Length 372;  
Best Local Similarity 78.8%; Pred. No. 1.1e-10;  
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 33  
DB 160 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 192

## RESULT 8

A35046  
E-selectin precursor - human  
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1); LECAM-2  
C:Species: Homo sapiens (man)  
C>Date: 27-Jul-1990 #sequence\_revision 27-Jul-1990 #text\_change 21-Jul-2000  
C:Accession: A38615; A35046; A32606  
R:Collins, T.; Williams, A.; Johnston, G.I.; Kim, J.; Eddy, R.; Shows, T.; Gimbrone Jr., J. Biol. Chem. 266, 2466-2473, 1991  
A:Title: Structure and chromosomal location of the gene for endothelial-leukocyte adhesion molecule  
A:Reference number: A38615; MUID:91115870  
A:Accession: A38615  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-610 <COL>  
A:Cross-references: GB:M61893; GB:M58017; NID:9182043; PIDN:AAA52375.1; PID:9182046  
R:Hession, C.; Osborn, L.; Goff, D.; Chi-Rosso, G.; Vassallo, C.; Pasek, M.; Pittack, C. Proc. Natl. Acad. Sci. U.S.A. 87, 1673-1677, 1990  
A:Title: Endothelial leukocyte adhesion molecule 1: direct expression cloning and functional characterization  
A:Reference number: A35046; MUID:90175359  
A:Accession: A35046  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-610 <HES>  
A:Cross-references: GB:M30640; NID:9182047; PIDN:AAA52377.1; PID:9182048  
R:Bevilacqua, M.P.; Stangelin, S.; Gimbrone Jr., M.A.; Seed, B. Science 243, 1160-1165, 1989  
A:Title: Endothelial leukocyte adhesion molecule 1: an inducible receptor for neutrophils  
A:Reference number: A32606; MUID:89162047  
A:Accession: A32606  
A:Molecule type: mRNA  
A:Residues: 1-467, 'Y', 469-610 <BEV>  
A:Cross-references: GB:M24736; NID:9537523; PIDN:AAA52376.1; PID:9537524  
C:Genetics:  
A:Gene: GDB:SELE; ELAM; ESEL; ELAM1  
A:Cross-references: GDB:120612; OMIM:131210  
A:Map position: 1q22-1q25  
C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor  
C:Keywords: duplication; glycoprotein; tandem repeat; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:12-138/Domain: C-type lectin homology <LCH>  
F:22-610/Product: endothelial leukocyte adhesion molecule 1 #status predicted <MAT>  
F:143-174/Domain: EGF homology <EGF>  
F:180-237/Domain: complement factor H repeat homology <FH01>  
F:242-299/Domain: complement factor H repeat homology <FH02>  
F:304-362/Domain: complement factor H repeat homology <FH03>  
F:367-425/Domain: complement factor H repeat homology <FH04>  
F:430-488/Domain: complement factor H repeat homology <FH05>  
F:493-547/Domain: complement factor H repeat homology <FH06>  
F:557-578/Domain: transmembrane #status predicted <TM>  
F:575,145,160,179,199,203,265,312,332,503,527/Binding site: carbohydrate (Asn) (covalent)

Query Match 60.5%; Score 130; DB 2; Length 610;  
Best Local Similarity 63.6%; Pred. No. 3e-07;  
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 33  
DB 143 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 175

## RESULT 9

JN0473  
P-selectin precursor - bovine  
N:Alternate names: granule membrane protein-140  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 11-Jan-2000  
C:Accession: JN0473  
R:Strubel, N.A.; Nguyen, M.; Kansas, G.S.; Tedder, T.F.; Bischoff, J. Biochem. Biophys. Res. Commun. 192, 338-344, 1993  
A:Title: Isolation and characterization of a bovine cDNA encoding a functional homologue of human P-selectin  
A:Reference number: JN0473; MUID:93249394  
A:Accession: JN0473  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-646 <STR>  
A:Cross-references: GB:L12041; NID:9304246; PIDN:AAA30743.1; PID:9304247  
C:Comment: This protein is a Ca2+ dependent receptor for myeloid cells.  
C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;  
C:Keywords: cell adhesion; glycoprotein; phosphatidyl; phosphoprotein; transmembrane  
F:1-41/Domain: signal sequence #status predicted <SIG>  
F:42-646/Product: P-selectin #status predicted <MAT>  
F:163-194/Domain: EGF homology <EGF>  
F:200-257/Domain: complement factor H repeat homology <FH1>  
F:262-319/Domain: complement factor H repeat homology <FH2>  
F:324-381/Domain: complement factor H repeat homology <FH3>  
F:386-443/Domain: complement factor H repeat homology <FH4>  
F:458-515/Domain: complement factor H repeat homology <FH5>  
F:520-577/Domain: complement factor H repeat homology <FH6>  
F:588-611/Domain: transmembrane #status predicted <TM>  
F:612-646/Domain: intracellular #status predicted <CYT>

Query Match 59.1%; Score 127; DB 2; Length 646;  
Best Local Similarity 57.6%; Pred. No. 6.9e-07;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 33  
DB 163 CQPMSCSGHCEVCETIINNHTCNCVGYGPQCQ 195

## RESULT 10

B42755  
E-selectin precursor - mouse  
N:Alternate names: endothelial leukocyte adhesion molecule 1 (ELAM-1)  
C:Species: Mus musculus (house mouse)  
C>Date: 04-Mar-1993 #sequence\_revision 31-Dec-2000 #text\_change 31-Dec-2000  
C:Accession: S23174; B42755  
R:Becker Andre, M.; van Huijsduijnen, R.H.; Losberger, C.; Whelan, J.; Delamarier, J. Eur. J. Biochem. 206, 401-411, 1992  
A:Title: Murine endothelial leukocyte adhesion molecule 1 is a close structural and functional homologue of human E-selectin  
A:Reference number: S23174; MUID:92283265  
A:Accession: S23174  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-612 <BEC>  
A:Cross-references: GB:M80778; NID:9193014; PIDN:AAA37547.1; PID:9193015  
R:Weller, A.; Isenmann, S.; Vestweber, D. J. Biol. Chem. 267, 15176-15183, 1992  
A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-selectin in endothelial cells  
A:Reference number: B42755; MUID:92340571  
A:Accession: B42755  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 'MKATAGV', 1-389, 391-612 <WEL>  
A:Cross-references: GB:M87862; NID:9193107  
A:Experimental source: endothelial cells  
A:Note: Sequence extracted from NCBI backbone (NCBI:109470)  
A:Note: the sequence in GenBank entry MUSELELC, release 117.0, (PIDN:AAA37577.1; PID:9193015) is uncertain whether the initiator is Met-1 or the AUG codon preceding the start codon  
C:Keywords: glycoprotein; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:12-138/Domain: C-type lectin homology <LCH>

F:22-612/Product: P-selectin #status predicted <MAT>  
 F:143-174/Domain: EGF homology <EGF>  
 F:180-238/Domain: complement factor H repeat homology <FH1>  
 F:243-300/Domain: complement factor H repeat homology <FH2>  
 F:305-363/Domain: complement factor H repeat homology <FH3>  
 F:368-426/Domain: complement factor H repeat homology <FH4>  
 F:431-489/Domain: complement factor H repeat homology <FH5>  
 F:494-548/Domain: complement factor H repeat homology <FH6>  
 F:25-391,528/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 58.1%; Score 126; DB 2; Length 612;  
 Best Local Similarity 57.6%; Pred. No. 8-5e-07;  
 Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33  
 DB 143 CPMSCSGHCEIETINSYCKCHPGLGPNC 175

## RESULT 11

JC5092  
 E-selectin - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 31-Jan-1997 #sequence\_revision 31-Jan-1997 #text\_change 19-May-2000  
 R:Winkler, H.; Brostjan, C.; Csizmadia, V.; Natarajan, G.; Anrather, J.; Bach, F.H.  
 Gene 176, 67-72, 1996  
 A:Title: The intron-exon structure of the porcine E-selectin-encoding gene.  
 A:Reference number: JC5092; MUID:97075911  
 A:Contents: endothelial cells  
 A:Accession: JC5092

A:Molecule type: DNA  
 A:Residues: 1-482 <MIN>  
 A:Cross-references: GB:U07521; NID:g1052974; PIDN:AAC48680.1; PID:g1052975  
 C:Comment: This protein is a member of the selectin family of adhesion molecules.  
 C:Genetics:

A:Introns: 13/1: 1421; 178/1: 237/1: 300/1: 363/1: 422/1: 459/1: 466/1  
 C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement facto  
 F:13-139/Domain: C-type lectin homology <CH>  
 F:181-233/Domain: complement factor H repeat homology <FH1>  
 F:240-298/Domain: complement factor H repeat homology <FH2>  
 F:303-361/Domain: complement factor H repeat homology <FH3>  
 F:366-420/Domain: complement factor H repeat homology <FH4>

Query Match 58.1%; Score 125; DB 2; Length 482;  
 Best Local Similarity 60.6%; Pred. No. 9.1e-07;  
 Matches 20; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33  
 DB 144 CPMSCSGHCEIETINSYCKCHPGLGPNC 176

## RESULT 12

A30359  
 P-selectin precursor - human  
 N:Alternate names: CD62 antigen; granule membrane protein 140  
 C:Species: Homo sapiens (man)  
 C:Date: 18-Oct-1989 #sequence\_revision 30-Sep-1991 #text\_change 19-May-2000  
 C:Accession: A30359  
 R:Johnston, G.T.; Cook, R.G.; McEver, R.P.  
 Cell 56, 1033-1044, 1989

A:Title: Cloning of GMP-140, a granule membrane protein of platelets and endothelium; se  
 A:Reference number: A30359; MUID:89168432  
 A:Accession: A30359

A:Molecule type: mRNA  
 A:Residues: 1-830 <JOH>  
 A:Cross-references: GB:M25322  
 A:Note: parts of this sequence, including the amino end of the mature protein, were conf  
 C:Genetics:  
 A:Gene: GDB:SELP; GRMP

A:Cross-references: GDB:120018; OMIM:173610

A:Map position: 1922-1925

C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;

C:Keywords: cell adhesion; glycoprotein; phospholipid; phospholipid; surface an

F:1-41/Domain: signal sequence #status predicted <SIG>

F:42-830/Product: P-selectin #status experimental <MAT>

F:163-194/Domain: EGF homology <EGF>

F:200-257/Domain: complement factor H repeat homology <FH01>

F:262-319/Domain: complement factor H repeat homology <FH02>

F:324-381/Domain: complement factor H repeat homology <FH03>

F:386-443/Domain: complement factor H repeat homology <FH04>

F:448-505/Domain: complement factor H repeat homology <FH05>

F:510-567/Domain: complement factor H repeat homology <FH06>

F:572-629/Domain: complement factor H repeat homology <FH07>

F:642-699/Domain: complement factor H repeat homology <FH08>

F:704-761/Domain: complement factor H repeat homology <FH09>

F:772-795/Domain: transmembrane #status predicted <TMN>

F:796-830/Domain: intracellular #status predicted <CVT>

F:54,98,180,212,219,411,460,518,665,716,723,741/Binding site: carbohydrate (Asn) (cov

Query Match 58.1%; Score 125; DB 2; Length 830;  
 Best Local Similarity 57.6%; Pred. No. 1.4e-06;  
 Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHCEVEIINHTCNCNDVGYGPOCQ 33  
 DB 163 CQPMSCSGHCEIETINSYCKCHPGLGPNC 195

## RESULT 13

P-selectin precursor - mouse  
 N:Alternate names: CD62; granule membrane protein 140; PADSEM  
 C:Species: Mus musculus (house mouse)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 19-May-2000  
 C:Accession: A42755; A44899

R:Weller, A.; Isenmann, S.; Vestweber, D.  
 J. Biol. Chem. 267, 15176-15183, 1992

A:Title: Cloning of the mouse endothelial selectins. Expression of both E- and P-sele

A:Reference number: A42755; MUID:92340571

A:Accession: A42755

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-768 <WEI>

A:Cross-references: GB:M87861; NID:g200552; PIDN:AAA40008.1; PID:g200553

A:Experimental source: endothelial cells

A:Note: sequence extracted from NCBI backbone (NCBI:109467)

R:Sanders, W.E.; Wilton, R.W.; Ballantyne, C.M.; Beaudet, A.L.  
 Blood 80, 795-800, 1992

A:Title: Molecular cloning and analysis of in vivo expression of murine P-selectin.

A:Reference number: A44899; MUID:92345617

A:Accession: A44899

A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-723, 'E', 725-768 <SAN>

A:Cross-references: GB:M72332; NID:g193565; PIDN:AAA37712.1; PID:g193566

A:Note: sequence extracted from NCBI backbone (NCBI:109900)

C:Superfamily: unassigned EGF-related proteins; complement factor H repeat homology;

C:Keywords: cell adhesion; glycoprotein; phospholipid; phospholipid; transmembr

F:1-11/Domain: signal sequence #status predicted <SIG>

F:42-768/Product: P-selectin #status predicted <MAT>

F:163-194/Domain: EGF homology <EGF>

F:200-257/Domain: complement factor H repeat homology <FH01>

F:262-319/Domain: complement factor H repeat homology <FH02>

F:324-381/Domain: complement factor H repeat homology <FH03>

F:386-443/Domain: complement factor H repeat homology <FH04>

F:448-505/Domain: complement factor H repeat homology <FH05>

F:510-567/Domain: complement factor H repeat homology <FH06>

F:580-637/Domain: complement factor H repeat homology <FH07>

F:642-699/Domain: complement factor H repeat homology <FH08>

F:710-733/Domain: transmembrane #status predicted <TMN>

F:734-768/Domain: intracellular #status predicted <INT>

F:45,54,107,212,347,398,456,467,603,654,661,679/Binding site: carbohydrate (Asn) (covalent)

Query Match 57.7%; Score 124; DB 2; Length 768;  
Best Local Similarity 54.5%; Pred. No. 1.7e-06;  
Matches 18; Conservative 7; Mismatches 8; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 33  
Db 163 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 195

RESULT 14  
S36772

E-selectin - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Dec-1993 #sequence\_revision 27-Feb-1997 #text\_change 19-May-2000  
C:Accession: S36772

R:Nguyen, M.; Strudel, N.A.; Bischoff, J.

Nature 365, 267-269, 1993

A:Title: A role for sialyl Lewis-X/A glycoconjugates in capillary morphogenesis.  
A:Reference number: S36772; PMID:93382537

A:Accession: S36772

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-485 <NCU>

A:Cross-references: GB:L12039; NID:g402913; PIDN:AAA02991.1; PID:g402914

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor

F:13-139/Domain: C-type lectin homology <LCH>

F:181-237/Domain: complement factor H repeat homology <FH1>

F:242-289/Domain: complement factor H repeat homology <FH2>

F:304-362/Domain: complement factor H repeat homology <FH3>

F:367-421/Domain: complement factor H repeat homology <FH4>

Query Match 54.9%; Score 118; DB 2; Length 485;  
Best Local Similarity 54.5%; Pred. No. 5.6e-06;  
Matches 18; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 33  
Db 144 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 176

RESULT 15  
I46709

endothelial leukocyte adhesion molecule 1 - rabbit

C:Species: Oryctolagus cuniculus (domestic rabbit)

C:Date: 14-Feb-1997 #sequence\_revision 14-Feb-1997 #text\_change 19-May-2000

C:Accession: I46709; I46708

R:Larigan, J.D.; Tsang, T.C.; Rumberger, J.M.; Burns, D.R.

DNA Cell Biol. 11, 149-162, 1992

A:Title: Characterization of cDNA and genomic sequences encoding rabbit ELAM-1: Conserva

A:Reference number: I46708; PMID:92189729

A:Accession: I46709

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-551 <LAR>

A:Cross-references: GB:M91005; NID:g165006; PIDN:AAA1244.1; PID:g165007

A:Accession: I46708

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-307, 'T', 309-327, 'T', 329-490, 'A', 492-551 <LAR2>

A:Cross-references: GB:M91004; NID:g165004; PIDN:AAA1243.1; PID:g165005

C:Genetics:

A:Gene: ELAM1

A:Map position: 1q22-q25

A:Insertions: 13/1; 143/1; 179/1; 241/1; 303/1; 366/1; 429/1; 488/1; 525/1; 533/2

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement factor

C:Keywords: duplication; glycoprotein; tandem repeat

F:14-140/Domain: C-type lectin homology <LCH>

F:182-239/Domain: complement factor H repeat homology <FH1>

F:244-301/Domain: complement factor H repeat homology <FH2>

F:306-364/Domain: complement factor H repeat homology <FH3>  
F:369-427/Domain: complement factor H repeat homology <FH4>  
F:432-486/Domain: complement factor H repeat homology <FH5>  
F:32,45,201,314,321,466/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 54.0%; Score 116; DB 2; Length 551;  
Best Local Similarity 54.5%; Pred. No. 1e-05;  
Matches 18; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 33  
Db 145 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 177

RESULT 16  
I53821

E-selectin - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 19-May-2000

C:Accession: I53821

R:Anchamach, J.A.; Oliver, M.G.; Anderson, D.C.; Manning, A.M.

Gene 145, 251-255, 1994

A:Title: Cloning, sequence comparison and in vivo expression of the gene encoding rat

A:Reference number: I53821; PMID:94333817

A:Accession: I53821

A:Molecule type: mRNA

A:Residues: 1-768 <RES>

A:Cross-references: GB:L23088; NID:g349552; PIDN:AAA60325.1; PID:g349553

C:Superfamily: unassigned EGF-related proteins; C-type lectin homology; complement fa

F:32-158/Domain: C-type lectin homology <LCH>

F:163-194/Domain: EGF homology <EGF>

F:262-319/Domain: complement factor H repeat homology <FHR>

F:510-567/Domain: complement factor H repeat homology <FH06>

F:580-637/Domain: complement factor H repeat homology <FH07>

F:642-699/Domain: complement factor H repeat homology <FH08>

Query Match 50.7%; Score 109; DB 2; Length 768;  
Best Local Similarity 48.5%; Pred. No. 8.2e-05;  
Matches 16; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 33  
Db 163 CQPMSCSGHGECEVEIINNHTCNCVDVGYGPQCQ 195

RESULT 17  
S28941

coagulation factor XIIa (EC 3.4.21.38) - guinea pig (fragment)

N:Alternate names: Hageman factor

C:Species: Cavia porcellus (guinea pig)

C:Date: 25-Feb-1994 #sequence\_revision 03-Aug-1995 #text\_change 21-Jan-2000

C:Accession: S28941

R:Semba, U.; Yamamoto, T.; Kunisada, T.; Shibuya, Y.; Tanase, S.; Kambara, T.; Okabe,

Biochim. Biophys. Acta 1159, 113-121, 1992

A:Title: Primary structure of guinea-pig Hageman factor: sequence around the cleavage

A:Reference number: S28941; PMID:93003367

A:Accession: S28941

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-603 <SEM>

A:Cross-references: EMBL:X68615; NID:g49578; PIDN:CAA48600.1; PID:g49579

C:Superfamily: coagulation factor XII; EGF homology; fibronectin type I repeat homo

C:Keywords: hydrolase; serine proteinase

F:6-87/Domain: fibronectin type II repeat homology <IF2>

F:134-169/Domain: fibronectin type I repeat homology <FI1>

F:177-208/Domain: EGF homology <EGF>

F:216-294/Domain: kringle homology <KR>

F:359-597/Domain: trypsin homology <TR>

```
Query Match          44.2%; Score 95; DB 2; Length 603;
Best Local Similarity 44.1%; Pred. No. 0.0025;
Matches 15; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINHTNCMDVGYGPOCOL 34
DB 177 CQTNPCNLNGRCLTEVEGHHLCDCPMGYTGPFCDL 210

RESULT 18
T21339
hypotheetical protein F25D7.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T21339
R:McMurray, A.
submitted to the EMBL Data Library, August 1996
A:Reference number: Z19409
A:Accession: T21339
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-710 <WIL>
A:Cross-references: EMBL:Z78418; PIDN:CA801698.1; GSPDB:GNO0019; CESP:F25D7.5
A:Experimental source: clone F25D7
C:Genetics:
A:Gene: CESP:F25D7.5
A:Map position: 1
A:introns: 20/1; 56/2; 87/1; 137/3; 235/3; 274/2; 410/2; 473/2; 496/3

Query Match          44.2%; Score 95; DB 2; Length 710;
Best Local Similarity 38.2%; Pred. No. 0.0029;
Matches 13; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINHTNCMDVGYGPOCOL 34
DB 148 CPTTCNGHCKCYDVEDVKDCYWGVEGHECEV 181

RESULT 19
A35844
Notch protein - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 12-Oct-1990 #sequence_revision 12-Oct-1990 #text_change 13-Aug-1999
C:Accession: A35844
R:Cotman, C.; Harris, W.; Kintner, C.
Science 249, 1438-1441, 1990
A:Title: Notch, the Xenopus homolog of Drosophila notch.
A:Reference number: A35844; MUID:90385285
A:Accession: A35844
A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
A:Molecule type: mRNA
A:Residues: 1-2524 <COF>
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology
C:Keywords: transmembrane protein
F:146-177/Domain: EGF homology <EGX1>
F:184-215/Domain: EGF homology <EGF1>
F:222-254/Domain: EGF homology <EGF>
F:456-487/Domain: EGF homology <EGX2>
F:757-788/Domain: EGF homology <EGF3>
F:1025-1056/Domain: EGF homology <EGX3>
F:1924-1956/Domain: ankyrin repeat homology <AN1>
F:1957-1989/Domain: ankyrin repeat homology <AN2>
F:1991-2023/Domain: ankyrin repeat homology <AN3>
F:2024-2056/Domain: ankyrin repeat homology <AN4>
F:2057-2089/Domain: ankyrin repeat homology <AN5>

Query Match          43.3%; Score 93; DB 2; Length 2524;
Best Local Similarity 45.5%; Pred. No. 0.013;
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINHTNCMDVGYGPOCOL 33

DB 911 CQPNCHNGSCSDGINMFNCNCPAGFRGPKCE 943

RESULT 20
T31070
notch homolog - sea urchin (Lytechinus variegatus)
C:Species: Lytechinus variegatus (variegated urchin)
C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 31-Jan-2000
C:Accession: T31070
R:Sherwood, D.R.; McClay, D.R.
Development 124, 3363-3374, 1997
A:Title: Identification and localization of a sea urchin Notch homologue: Insights in
A:Reference number: Z20966; MUID:97454256
A:Accession: T31070
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-2531 <SHE>
A:Cross-references: EMBL:AF000634; NID:g2570350; PID:g2570351; PIDN:AA802088.1
C:Superfamily: notch protein; ankyrin repeat homology; EGF homology

Query Match          41.9%; Score 90; DB 2; Length 2531;
Best Local Similarity 38.2%; Pred. No. 0.029;
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINHTNCMDVGYGPOCOL 34
DB 555 CQSPCENGTCIDGVOFTCLCTGVEGHRCEM 588

RESULT 21
S42373
hypotheetical protein T2065.3 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999
C:Accession: S42373
R:Smith, A.
submitted to the EMBL Data Library, March 1994
A:Reference number: S42368
A:Accession: S42373
A:Molecule type: DNA
A:Residues: 1-3051 <SMT>
A:Cross-references: EMBL:Z30423; NID:g458479; PID:g458485
C:Genetics:
A:introns: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/1; 952/2; 1041/1; 1146
C:Superfamily: von Willebrand factor type A repeat homology; EGF homology; fibronecti
F:512-679/Domain: von Willebrand factor type A repeat homology <VMA1>
F:754-793/Domain: fibronectin type II repeat homology <2FI>
F:1201-1244/Domain: EGF homology <EGF>

Query Match          41.9%; Score 90; DB 2; Length 3051;
Best Local Similarity 51.6%; Pred. No. 0.034;
Matches 16; Conservative 4; Mismatches 9; Indels 2; Gaps 1;

OY 6 CSGHGEVVE--IINHTNCMDVGYGPOCOL 34
DB 2664 CNAHGDVCHNFTATNNTTCVCTDGTGWTGPOCOV 2694

RESULT 22
A49175
Notch B protein - mouse (fragment)
N:Alternate names: Notch homolog
C:Species: Mus musculus (house mouse)
C:Date: 21-Jan-1994 #sequence_revision 05-Jan-1996 #text_change 20-Sep-1999
C:Accession: A49175; PH1570; S32113
R:Lardelli, M.; Lendahl, U.
Exp. Cell Res. 204, 364-372, 1993
A:Title: Notch A and Notch B--two mouse Notch homologues coexpressed in a wide variety
A:Reference number: A49175; MUID:93178563
A:Accession: A49175
```

A>Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-1203 <LARP>  
A:Cross-references: EMBL:X68279; NID:g287989; PIDN:CAA8340.1; PID:g287990  
A:Experimental source: embryo  
A>Note: sequence extracted from NCBI backbone (NCBIP:126158)  
C:Comment: This protein has many EGF repeats and 11n-12/Notch repeats.  
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology  
F:143-174/Domain: EGF homology <EGX1>  
F:482-513/Domain: EGF homology <EGF1>  
F:560-591/Domain: EGF homology <EGF>  
F:674-705/Domain: EGF homology <EGX2>  
F:712-743/Domain: EGF homology <EGF3>  
F:836-867/Domain: EGF homology <EGX3>

Query Match 41.4%; Score 89; DB 2; Length 1203;  
Best Local Similarity 41.2%; Pred. No. 0.021;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGECVELINNHNTNCNDVGYGPOCQ 34  
Db 181 CQSNPCVNNQCVDKVNRFQCLCPGFTGPVCQI 214

RESULT 23  
78549  
notch3 protein - human

C:Species: Homo sapiens (man)

C:Date: 24-Jul-1998 #sequence\_revision 24-Jul-1998 #text\_change 20-Sep-1999

C:Accession: S78549; S71825

R:Joutel, A.; Tournier-Lasserre, E.  
submitted to the EMBL Data Library, April 1997

A:Reference number: S78549

A:Accession: S78549

A:Molecule type: mRNA

A:Residues: 1-2321 <JOU1>

A:Cross-references: EMBL:097669; NID:g2668591; PIDN:AAB91371.1; PID:g2668592

X:Joutel, A.; Corpechot, C.; Ducros, A.; Vahedi, K.; Chabrier, H.; Mouton, P.; Alamowicz, M.M.; Weissbach, J.; Bach, J.F.; Bousser, M.G.; Tournier-Lasserre, E.

Nature 383, 707-710, 1996

A:Title: Notch3 mutations in CADASIL, a hereditary adult-onset condition causing stroke

A:Reference number: S71825; MUID:97032728

A:Accession: S71825

A>Status: nucleic acid sequence not shown

A:Molecule type: DNA

A:Residues: 67-113;138-194;268-333;335-346;536-613;716-765;1240-1279;1815-1888 <JOU2>

A:Cross-references: EMBL:097669

C:Genetics:

A:Gene: notch3

A:Map position: 19p13.1

C:Function:

A:Description: may be involved in pathogenesis of CADASIL, causing a type of stroke and

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homology

C:Keywords: tandem repeat; transmembrane protein

F:123-155/Domain: EGF homology <EGX1>

F:162-194/Domain: EGF homology <EGX1>

F:240-271/Domain: EGF homology <EGX2>

F:318-349/Domain: EGF homology <EGF>

F:473-504/Domain: EGF homology <EGX3>

F:553-584/Domain: EGF homology <EGF3>

F:928-959/Domain: EGF homology <EGX4>

F:1070-1126/Domain: laminin-type EGF-like homology <LEG>

F:1838-1870/Domain: ankyrin repeat homology <AN1>

F:1871-1903/Domain: ankyrin repeat homology <AN2>

F:1905-1937/Domain: ankyrin repeat homology <AN3>

F:1938-1970/Domain: ankyrin repeat homology <AN4>

F:1971-2003/Domain: ankyrin repeat homology <AN5>

Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;  
Qy 1 CQPMSCSGHGECVELINNHNTNCNDVGYGPOCQ 33  
Db 928 CSPSCFNGGTGCVDNVSFSCLCRPGYGAHCQ 960

RESULT 24  
A49128

cell fate determining gene Notch2 protein - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Jan-1994 #sequence\_revision 18-Nov-1994 #text\_change 13-Aug-1999

C:Accession: A49128

R:Weinmaster, G.; Roberts, V.J.; Lemke, G.  
Development 116, 931-941, 1992

A:Title: Notch2: a second mammalian Notch gene.

A:Reference number: A49128; MUID:93202015

A:Accession: A49128

A>Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-2471 <ME1>

A:Experimental source: Schwann cell

A>Note: sequence extracted from NCBI backbone (NCBIP:127811)

C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol

F:264-295/Domain: EGF homology <EGX1>

F:799-830/Domain: EGF homology <EGX1>

F:877-908/Domain: EGF homology <EGX2>

F:1029-1060/Domain: EGF homology <EGF>

F:1067-1098/Domain: EGF homology <EGX3>

F:1153-1184/Domain: EGF homology <EGF3>

F:1191-1222/Domain: EGF homology <EGX4>

F:1876-1908/Domain: ankyrin repeat homology <AN1>

F:1909-1941/Domain: ankyrin repeat homology <AN2>

F:1943-1975/Domain: ankyrin repeat homology <AN3>

F:1976-2008/Domain: ankyrin repeat homology <AN4>

F:2009-2041/Domain: ankyrin repeat homology <AN5>

Query Match 41.4%; Score 89; DB 2; Length 2471;  
Best Local Similarity 41.2%; Pred. No. 0.037;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGECVELINNHNTNCNDVGYGPOCQ 34  
Db 498 CQSNPCVNNQCVDKVNRFQCLCPGFTGPVCQI 531

RESULT 25  
JX0210

protein C (activated) (EC 3.4.21.69) precursor - mouse

N:Alternate names: vitamin K-dependent serine protease

C:Species: Mus musculus (house mouse)

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C:Accession: JX0210

R:Rade, N.; Sato, M.; Tsujimura, A.; Iwase, R.; Hashimoto-Gotoh, T.  
J. Biochem. 111, 431-495, 1992

A:Title: Isolation and characterization of a mouse protein C cDNA.

A:Reference number: JX0210; MUID:92316897

A:Accession: JX0210

A:Molecule type: mRNA

A:Residues: 1-461 <TAB>

A:Cross-references: GB:D10445; NID:g220385; PIDN:BAA01235.1; PID:g220386

A:Experimental source: liver

C:Comment: Protein C is the zymogen of the vitamin K-dependent serine proteinase that

C:Superfamily: coagulation factor X; EGF homology; Gla domain homology; trypsin homol

C:Keywords: beta-hydroxyaspartic acid; blood coagulation; calcium binding; carboxyglu

F:1-35/Domain: signal sequence #status predicted <SIG>

F:37-85/Domain: Gla domain homology <Gla>

F:34-41/Domain: propeptide #status predicted <PRO>

F:42-196,199-461/Product: protein C #status predicted <PRC>

F:42-196/Domain: light chain #status predicted <PCL>

F:91-130/Domain: EGF homology <EG1>



F:139-174/Domain: EGF homology <EG2>  
F:199-461/Domain: heavy chain #status predicted <PCH>  
F:199-219/Domain: activation peptide #status predicted <ACT>  
F:212-461/Product: vitamin K dependent serine proteinase #status predicted <VIT>  
F:212-443/Domain: trypsin homology <TRY>  
F:47-48,55,60,61,66,67,70,76/Modified site: gamma-carboxyglutamic acid (Glu) #status  
F:121-130,139-150,146-159,161-174,182-319,238-254,373-387,398-426/Disulfide bonds: #stat  
F:214,290,335/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:253,299,402/Active site: His, Asp, Ser #status predicted

Query Match 40.5%; Score 87; DB 1; Length 461;  
Best Local Similarity 46.4%; Pred. No. 0.016;

Matches 13; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 6 CSHGECVEIINNHTCNCDDVGYGPOCQ 33  
DB 104 CCGHGTICIDIGISFCSCDKGMEGKFCQ 131

RESULT 26

150719 C-Delta-1 - chicken

C:Species: Gallus gallus (chicken)

C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000

C:Accession: I50719

R:Hemique, D.; Adam, J.; Myat, A.; Chltnis, A.; Lewis, J.; Ish-Horowicz, D.

Nature 375, 787-790, 1995

A:Title: Expression of a Delta homologue in prospective neurons in the chick.

A:Reference number: I50719; MUID:95319507

A:Accession: I50719

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-728 <HN>

A:Cross-references: EMBL:026590; NID:g882411; PIDN:AAC59689.1; PID:g882412

C:Superfamily: unassigned EGF-related proteins; EGF homology

F:299-333/Domain: EGF homology <EGX1>

F:339-370/Domain: EGF homology <EGF1>

F:416-447/Domain: EGF homology <EGX2>

F:454-485/Domain: EGF homology <EGF>

F:492-523/Domain: EGF homology <EGF3>

Query Match 40.5%; Score 87; DB 2; Length 728;  
Best Local Similarity 35.3%; Pred. No. 0.023;

Matches 12; Conservative 7; Mismatches 15; Indels 0; Gaps 0;

QY 1 COPWSCGHCEVEIINNHTCNCDDVGYGPOCQ 34  
DB 339 CDANPCKNGSCTDLENSTCTCPGFYKNCBL 372

RESULT 27

A40136 fibropellin Ia - sea urchin (Strongylocentrotus purpuratus)

N:Alternate names: epidermal growth factor homolog precursor

M:Contains: alternatively spliced fibropellin Ib (EGF1)

C:Species: Strongylocentrotus purpuratus (purple urchin)

C:Date: 13-May-1992 #sequence\_revision 17-Sep-1997 #text\_change 21-Jul-2000

C:Accession: A40136; A40136; A29316; A43131

R:Delgadillo-Reynoso, M.G.; Rollo, D.R.; Hursh, D.A.; Raff, R.A.

J Mol. Evol. 29, 314-327, 1989

A:Title: Structural analysis of the uEGF gene in the sea urchin Strongylocentrotus purp

A:Reference number: A40136; MUID:90112459

A:Accession: A40136

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-114 <DEL>

A:Cross-references: GB:X17530; NID:g10225; PID:g667061

A:Accession: B40136

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 181-251,329-370,'R',372-408,'RA',411-441 <DE2>

A:Accession: C40136

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 'K',747-821,898-978 <DE3>

R:Hursh, D.A.; Andrews, M.E.; Raff, R.A.

Science 237, 1487-1490, 1987

A:Title: A sea urchin gene encodes a polypeptide homologous to epidermal growth facto

A:Reference number: A29316; MUID:87319677

A:Accession: A29316

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 'S',280-481,786-1064 <HUR>

A:Cross-references: GB:M17421; NID:g161474; PIDN:AAA30050.1; PID:g552260

R:Hunt, L.T.; Barker, W.C

FASEB J. 3, 1760-1764, 1989

A:Title: Avidin-like domain in an epidermal growth factor homolog from a sea urchin.

A:Reference number: A43131; MUID:89196806

A:Contents: annotation

C:Comment: EGF homology repeats 10-17 are spliced out in the short form (fibropellin

F:1-19/Domain: signal sequence #status predicted <SIG>

F:20-1064/Product: fibropellin I #status predicted <PIB>

F:23-54/Domain: EGF homology <EG01>

F:57-175/Domain: C1r/C1s repeat homology <C1R>

F:180-211/Domain: EGF homology <EG02>

F:218-249/Domain: EGF homology <EG03>

F:256-287/Domain: EGF homology <EG04>

F:324-325/Domain: EGF homology <EG05>

F:332-363/Domain: EGF homology <EG06>

F:370-401/Domain: EGF homology <EG07>

F:408-439/Domain: EGF homology <EG08>

F:446-477/Domain: EGF homology <EG09>

F:484-515/Domain: EGF homology <EG10>

F:522-553/Domain: EGF homology <EG11>

F:560-591/Domain: EGF homology <EG12>

F:598-629/Domain: EGF homology <EG13>

F:636-667/Domain: EGF homology <EG14>

F:674-705/Domain: EGF homology <EG15>

F:712-743/Domain: EGF homology <EG16>

F:750-781/Domain: EGF homology <EG17>

F:788-819/Domain: EGF homology <EG18>

F:826-857/Domain: EGF homology <EG19>

F:864-895/Domain: EGF homology <EG20>

F:902-933/Domain: EGF homology <EG21>

F:936-1064/Region: avidin-like

F:23-34,28-43,45-54,62-88,180-191,185-200,202-211,218-229,223-238,240-249,256-267,261

57,451-466,468-477,484-495/Disulfide bonds: #status predicted

F:489-504,506-515,522-533,527-543,544-553,560-571,565-580,582-591,598-609,603-618,620

06,810-819,826-837,831-846,848-857,864-875,869-884,886-895,902-913,907-922,924-933/DI

Query Match 40.5%; Score 87; DB 2; Length 1064;  
Best Local Similarity 39.4%; Pred. No. 0.031;

Matches 13; Conservative 5; Mismatches 15; Indels 0; Gaps 0;

QY 1 COPWSCGHCEVEIINNHTCNCDDVGYGPOCQ 33  
DB 750 CASMPINGACIEWNGVTCOCVAGYGVCE 782

RESULT 28

A35672 crumbs protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 21-Sep-1990 #sequence\_revision 18-Nov-1992 #text\_change 11-Jan-2000

C:Accession: A35672

R:Teppas, U.; Theres, C.; Knust, E.

Cell 61, 787-799, 1990

A:Title: crumbs encodes an EGF-like protein expressed on apical membranes of Drosophi

A:Reference number: A35672; MUID:90263104

A:Accession: A35672

A:Status: preliminary



A:Cross-references: EMBL:U24703; NID:g902486; PID:g902487  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
F:1-27/Domain: signal sequence #status predicted <Sig>  
F:28-3461/Product: reelin #status predicted <MAT>  
F:1769-1795/Domain: EGF homology <EGF>

Query Match 39.8%; Score 85.5; DB 2; Length 3461;  
Best Local Similarity 51.7%; Pred. No. 0.12;  
Matches 15; Conservative 2; Mismatches 9; Indels 3; Gaps 1;

QY 6 CSGHGECEIINNHNTCNCVGYGPOCQL 34  
DB 2137 CYGHGSC---INGTKICIDPGYSGPTCKI 2162

RESULT 32  
EB9753  
protein F11C7.4 [Imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 09-Nov-2001  
C:Accession: EB9753  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A75000; MUID:99069613; PMID:9851916  
A:Note: see websites genome.wustl.edu/gsc/C\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and  
A:Accession: EB9753  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1722 <STO>  
A:Cross-references: GB:chr\_X; PIDN:AAC69012.1; PID:g1125776; GSPDB:GN00028; CESP:F11C7.4  
C:Genetics:  
A:Gene: F11C7.4  
A:Map position: X

Query Match 39.5%; Score 85; DB 2; Length 1722;  
Best Local Similarity 42.4%; Pred. No. 0.077;  
Matches 14; Conservative 3; Mismatches 16; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHNTCNCVGYGPOCQ 33  
DB 1371 CSSNTCSSRGACSPVWNNTVYCNCDNMWRGAHQ 1403

RESULT 33  
T30201  
Notch homolog protein - sea squirt (Halocynthia roretzi)  
C:Species: Halocynthia roretzi  
C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 02-Sep-2000  
C:Accession: T30201  
R:Hoti, S.; Saitoh, T.; Matsumoto, M.; Makabe, K.W.; Nishida, H.  
Dev. Genes Evol. 207, 371-380, 1997  
A:Title: Notch homologue from Halocynthia roretzi is preferentially expressed in the cen  
A:Reference number: 220775  
A:Accession: T30201  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-2352 <HOR>  
A:Cross-references: EMBL:AB001327; NID:d1204472; PID:d1026501; PIDN:BA25571.1  
C:Genetics:  
A:Gene: Notch

Query Match 39.5%; Score 85; DB 2; Length 2352;  
Best Local Similarity 43.8%; Pred. No. 0.1;  
Matches 14; Conservative 2; Mismatches 16; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINNHNTCNCVGYGPOC 32  
DB 113 CSPNCSNGACEELNSFKCTGSGYVDTC 144

RESULT 34  
T29359  
hypothetical protein R05G6.9 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T29359  
R:Murray, J.; Le, T.T.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid R05G6.  
A:Reference number: 220612  
A:Accession: T29359  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-372 <MUR>  
A:Cross-references: EMBL:U58746; PIDN:AAB00626.1; GSPDB:GN00022; CESP:R05G6.9  
A:Experimental source: strain Bristol N2; clone R05G6  
C:Genetics:  
A:Gene: CESP:R05G6.9  
A:Map position: 4  
A:introns: 80/1; 161/1; 245/1; 286/1

Query Match 39.1%; Score 84; DB 2; Length 372;  
Best Local Similarity 40.7%; Pred. No. 0.029;  
Matches 11; Conservative 9; Mismatches 7; Indels 0; Gaps 0;

QY 6 CSGHGECEIINNHNTCNCVGYGPOC 32  
DB 339 CSGNGICISLKSYSCTCNLGMTGPTC 365

RESULT 35  
JQ1322  
tenascin precursor - mouse  
N:Alternate names: contractin; hexabrachion  
C:Species: Mus musculus (house mouse)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: JQ1322; A37936; B37936; S14571; S50209  
R:Sage, Y.; Tsukamoto, T.; Jing, N.; Kusakabe, M.; Sakakura, T.  
Gene 104, 177-185, 1991  
A:Title: Murine tenascin: cDNA cloning, structure and temporal expression of isoforms  
A:Reference number: JQ1322; MUID:92009211  
A:Accession: JQ1322  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-2019 <SAG>  
A:Cross-references: GB:D90343; NID:g220609; PIDN:BA14355.1; PID:g220610  
A:Experimental source: cell line 2H6GR  
A:Note: the authors translated the codon ATG for residue 60 as Trp  
R:Weller, A.; Beck, S.; Ekblom, P.  
J. Cell Biol. 112, 355-362, 1991  
A:Title: Amino acid sequence of mouse tenascin and differential expression of two ten  
A:Reference number: A37936; MUID:91107734  
A:Accession: A37936  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 10  
A:Cross-references: GB:X56304  
A:Accession: B37936  
A:Status: preliminary; nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-620, 622-1010, 'N', 1012-1018, 'S', 1020-1024, 'H', 10  
A:Cross-references: GB:X56304  
R:Weller, A.; Beck, S.; Ekblom, P.  
submitted to the EMBL Data Library, August 1990  
A:Description: Amino acid sequence of mouse tenascin and differential expression of tw  
A:Reference number: S14571  
A:Accession: S14571  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-201, 'E', 203-317, 'S', 319-1018, 'S', 1020-1024, 'H', 1026-1305, 'S', 1307-2019

A:Cross-references: EMBL:X56304; NID:954768; PIDN:CAA39751.1; PID:954769  
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.  
Biochim. Biophys. Acta 1219, 613-622, 1994  
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues by  
A:Reference number: S50206; MUID:95035091  
A:Accession: S50209  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 46-146 <GLU>  
A:Cross-references: EMBL:X80281  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin type  
C:Keywords: alternative splicing; calcium binding; cell adhesion; duplication; extracellular  
F:1-22/Domain: signal sequence #status predicted <SIG>  
F:23-2019/Product: tenascin, long splice form #status predicted <MAT>  
F:23-1071,1527-2019/Product: tenascin, short splice form #status predicted <MAT2>  
F:408-434/Domain: EGF homology <EGF>  
F:622-703/Domain: fibronectin type III repeat homology <FN3A>  
F:711-793/Domain: fibronectin type III repeat homology <FN3B>  
F:802-884/Domain: fibronectin type III repeat homology <FN3C>  
F:892-976/Domain: fibronectin type III repeat homology <FN3D>  
F:984-1064/Domain: fibronectin type III repeat homology <FN3E>  
F:1073-1155/Domain: fibronectin type III repeat homology <FN3F>  
F:1346-1428/Domain: fibronectin type III repeat homology <FN3G>  
F:1437-1519/Domain: fibronectin type III repeat homology <FN3H>  
F:1527-1608/Domain: fibronectin type III repeat homology <FN3I>  
F:1616-1696/Domain: fibronectin type III repeat homology <FN3J>  
F:1704-1784/Domain: fibronectin type III repeat homology <FN3K>  
F:1799-2007/Domain: fibronectin beta/gamma homology <FBG>  
F:38,166,184,327,788,1018,1079,1093,1119,1184,1210,1275,1301,1354,1364,1394,1627,1878,19  
Query Match 39.1%; Score 84; DB 1; Length 2019;  
Best Local Similarity 53.3%; Pred. No. 0.11;  
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;  
OY 3 PWSGSGHGEVEIINHTCNCNVGYGPGC 32  
Db 409 PNGCSGHRGV----NGQCVCDEGTGEDC 434  
RESULT 36  
A32160  
tenascin-C - human  
N:Alternate names: hexabrachion  
C:Species: Homo sapiens (man)  
C:Date: 31-Jul-1989 #sequence.revision 12-Apr-1996 #text.change 20-Aug-1999  
C:Accession: I38337; A32160; S14015; S16166; S50208; S49354  
R:Gherzi, R.; Carnemolla, B.; Sirtl, A.; Fonassi, M.; Balza, E.; Zardi, L.  
J. Biol. Chem. 270, 3429-3434, 1995  
A:Title: Human tenascin gene. Structure of the 5'-region, identification, and character  
A:Reference number: A55974; MUID:95155442  
A:Accession: I38337  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2201 <RES>  
A:Cross-references: EMBL:X78565; NID:9556844; PIDN:CAA55309.1; PID:9556845  
R:Gulcher, J.R.; Nies, D.E.; Marton, L.S.; Stefansson, K.  
Proc. Natl. Acad. Sci. U.S.A. 86, 1588-1592, 1989  
A:Title: An alternatively spliced region of the human hexabrachion contains a repeat of  
A:Reference number: A32160; MUID:85160821  
A:Accession: A32160  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 431-538, 'R', 540-1065, 'H', 1067-1599, 'LMLHPRASN', 1609-2054, 'L' <GUL>  
A:Cross-references: GB:M24630; NID:9514363; PIDN:AAA52703.1; PID:9553348  
R:Sirtl, A.; Carnemolla, B.; Seginani, M.; Leprini, A.; Casari, G.; Baralle, F.; Zardi, L.  
Nucleic Acids Res. 19, 525-531, 1991  
A:Title: Human tenascin: primary structure, pre-mRNA splicing patterns and localization  
A:Reference number: S14015; MUID:91187670  
A:Accession: S14015  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 1-243,245-369, 'V', 371-679, 'R', 681-1676, 'I', 1678-2139, 'TR', 2142,2144-2201  
A:Cross-references: EMBL:X56160; NID:937226; PIDN:CAA39628.1; PID:937227  
R:Nies, D.E.; Hemeseth, T.J.; Kim, J.H.; Gulcher, J.R.; Stefansson, K.  
J. Biol. Chem. 266, 2818-2823, 1991  
A:Title: The complete cDNA sequence of human hexabrachion (tenascin). A multidomain p  
A:Reference number: S16166; MUID:91131572  
A:Accession: S16166  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-538, 'R', 540-1065, 'H', 1067-1599, 'LMLHPRASN', 1609-2054, 'LH', 2055-2201 <NI  
R:Glumoff, V.; Savontaus, M.; Vehanen, J.; Vuorio, E.  
Biochim. Biophys. Acta 1219, 613-622, 1994  
A:Title: Analysis of aggrecan and tenascin gene expression in mouse skeletal tissues  
A:Reference number: S50206; MUID:95035091  
A:Accession: S50208  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: mRNA  
A:Residues: 46-125 <GLU>  
A:Cross-references: EMBL:X80280  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, July 1994  
C:Genetics:  
A:Gene: GDB:HXB  
A:Cross-references: GDB:120073; OMIM:187380  
A:Map position: 9q33-q9q33  
C:Superfamily: tenascin; EGF homology; fibrinogen beta/gamma homology; fibronectin ty  
C:Keywords: alternative splicing; extracellular matrix  
F:622-703/Domain: EGF homology <EGF>  
F:408-434/Domain: fibronectin type III repeat homology <FN3A>  
F:711-794/Domain: fibronectin type III repeat homology <FN3B>  
F:892-976/Domain: fibronectin type III repeat homology <FN3C>  
F:984-1064/Domain: fibronectin type III repeat homology <FN3D>  
F:1437-1519/Domain: fibronectin type III repeat homology <FN3E>  
F:1619-1701/Domain: fibronectin type III repeat homology <FN3F>  
F:1709-1790/Domain: fibronectin type III repeat homology <FN3G>  
F:1798-1878/Domain: fibronectin type III repeat homology <FN3H>  
F:1886-1966/Domain: fibronectin type III repeat homology <FN3I>  
F:1981-2189/Domain: fibrinogen beta/gamma homology <FBG>  
Query Match 39.1%; Score 84; DB 2; Length 2201;  
Best Local Similarity 53.3%; Pred. No. 0.12;  
Matches 16; Conservative 0; Mismatches 10; Indels 4; Gaps 1;  
OY 3 PWSGSGHGEVEIINHTCNCNVGYGPGC 32  
Db 409 PNGCSGHRGV----NGQCVCDEGTGEDC 434  
RESULT 37  
S45306  
notch 3 protein - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Feb-1995 #sequence.revision 20-Feb-1995 #text.change 20-Sep-1999  
C:Accession: S45306  
R:Liardelli, M.; Dahlstrand, J.; Lendahl, U.  
Mech. Dev. 46, 123-136, 1994  
A:Title: The novel Notch homologue mouse Notch 3 lacks specific epidermal growth fact  
A:Reference number: S45306; MUID:95001556  
A:Accession: S45306  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-2318 <LAR>  
A:Cross-references: EMBL:X74760; NID:9483580; PIDN:CAA52276.1; PID:9483581  
C:Superfamily: unassigned ankyrin repeat proteins; ankyrin repeat homology; EGF homol  
F:163-195/Domain: EGF homology <EGF1>  
F:474-505/Domain: EGF homology <EGF2>  
F:854-885/Domain: EGF homology <EGF3>  
F:1839-1871/Domain: ankyrin repeat homology <AN1>  
F:1872-1904/Domain: ankyrin repeat homology <AN2>  
F:1906-1938/Domain: ankyrin repeat homology <AN3>  
F:1939-1971/Domain: ankyrin repeat homology <AN4>  
F:1972-2004/Domain: ankyrin repeat homology <AN5>

Query Match 39.1%; Score 84; DB 2; Length 2318;  
Best Local Similarity 42.4%; Pred. No. 0.13;  
Matches 14; Conservative 5; Mismatches 14; Indels 0; Gaps 0;

OY 1 COPWSCGHGECVEIINHTCNCVDVGYGPOC 33  
DB 929 CSPSSCFNGCTCVDGVSFSCLCRPGYTGTTCQ 961

## RESULT 38

T09456  
Intrinsic factor-B12 receptor Cubillin precursor - human

C:Species: Homo sapiens (man)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: T09456  
R:Koziyal, R.; Kristiansen, M.; Sllahtaroglu, A.; Hansen, C.; Jacobsen, C.; Tommerup, N.  
Blood 91, 3593-3600, 1998

A:Title: The human intrinsic factor-vitamin B12 receptor, cubillin: Molecular characterization.  
A:Reference number: Z16677; MUID:98241400  
A:Accession: T09456  
A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-3623 <KOZ>  
A:Cross-references: EMBL:AF034611; NID:g3929528; PIDN:AAC82612.1; PID:g3929529  
C:Genetics:

A:Map position: 10p12  
C:Superfamily: unassigned EGF-related proteins; EGF homology  
C:Keywords: receptor; vitamin B12 uptake

F:1-24/Domain: signal sequence #status predicted <SIG>  
F:25-3623/Product: intrinsic factor-B12 receptor #status predicted <MAT>  
F:436-467/Domain: EGF homology <EGF>

Query Match 39.1%; Score 84; DB 2; Length 3623;  
Best Local Similarity 43.3%; Pred. No. 0.18;  
Matches 13; Conservative 5; Mismatches 12; Indels 0; Gaps 0;

OY 3 PWSGSGHGECEIINHTCNCVDVGYGPOC 32  
DB 181 PLSCONGCTCVNTMGSYCHCPETGYGPOC 210

## RESULT 39

T42681

hypothetical protein DKFZp434E0321.1 - human (fragments)  
C:Species: Homo sapiens (man)  
C:Date: 11-Jan-2000 #sequence\_revision 11-Jan-2000 #text\_change 11-Jan-2000

C:Accession: T42681  
R:Blum, H.; Bauerach, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, November 1999

A:Reference number: Z22233  
A:Accession: T42681  
A:Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-870; 871-1069 <AAA>  
A:Cross-references: EMBL:AL133021

A:Experimental source: adult testis; clone DKFZp434E0321  
A:Note: the cDNA sequence contains a +1 frameshift near codon 870  
C:Genetics:  
A:Note: DKFZp434E0321.1

Query Match 38.8%; Score 83.5; DB 2; Length 1069;  
Best Local Similarity 42.4%; Pred. No. 0.078;  
Matches 14; Conservative 5; Mismatches 13; Indels 1; Gaps 1;

OY 1 COPWSCGHGECVEIINHTCNCVDVGYGPOC 32  
DB 563 CLPGCSIDHGQCDGDTGSGQCLCETGWTGPSC 595

## RESULT 40

T13953

MEGF5 protein - rat

N:Alternate names: slit protein homolog  
C:Species: Rattus norvegicus (Norway rat)

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T13953  
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998

A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: Z14126; MUID:98360089  
A:Accession: T13953

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA

A:Residues: 1-1523 <NAK>  
A:Cross-references: EMBL:AB011531; NID:g3449291; PIDN:BAA32461.1; PID:g3449292  
C:Genetics:  
A:Gene: MEGF5

Query Match 38.8%; Score 83.5; DB 2; Length 1523;  
Best Local Similarity 50.0%; Pred. No. 0.1;  
Matches 16; Conservative 2; Mismatches 13; Indels 1; Gaps 1;

OY 1 COPWSCGHGECVEIINHTCNCVDVGYGPOC 32  
DB 1372 CLGHSCS-HGTCVANGNSYVCKACBEGSPIC 1402

Search completed: September 7, 2002, 10:15:28  
Job time: 206 sec

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:22:08 : Search time 85.16 Seconds  
(without alignments)  
100.974 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193  
Perfect score: 215  
Sequence: 1 COPWSCSGHGECEIINNHTCNCVDVGYGPOCOL 34

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending\_Patents\_AA\_New:\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	372	5	US-09-791-537-152667
2	209	97.2	372	5	US-09-791-537-84593
3	205	95.3	323	5	US-09-791-537-53485
4	205	95.3	341	6	US-10-211-364-1171
5	205	95.3	341	6	US-10-212-054-1328
6	205	95.3	341	6	US-10-212-778-1158
7	205	95.3	363	5	US-09-791-537-58446
8	205	95.3	372	5	US-09-791-537-22816
9	205	95.3	372	5	US-09-791-537-42655
10	205	95.3	372	5	US-09-791-537-42657
11	205	95.3	372	5	US-09-791-537-42658
12	205	95.3	372	5	US-09-791-537-42659
13	205	95.3	372	5	US-09-791-537-69658
14	205	95.3	385	5	US-09-791-537-51391
15	205	95.3	385	5	US-09-791-537-53844
16	205	95.3	385	5	US-09-791-537-113060
17	191	88.8	370	5	US-09-791-537-50403
18	181	84.2	376	5	US-09-791-537-132144
19	169	78.6	360	5	US-09-791-537-60503
20	159	74.0	372	5	US-09-791-537-37750
21	159	74.0	372	5	US-09-791-537-6693
22	159	74.0	372	5	US-09-791-537-81233
23	131	60.9	611	5	US-09-791-537-118878
24	130	60.5	162	5	US-09-791-537-23269
25	130	60.5	610	1	PCR-US02-23913-357
26	130	60.5	610	5	US-09-791-537-22813

27	130	60.5	610	5	US-09-791-537-121834	Sequence 121834,
28	130	60.5	610	6	US-10-205-823-357	Sequence 357, App
29	127	59.1	549	5	US-09-791-537-50406	Sequence 50406, A
30	127	59.1	646	5	US-09-791-537-84829	Sequence 84829, A
31	127	59.1	649	5	US-09-791-537-93873	Sequence 93873, A
32	127	59.1	769	5	US-09-791-537-50409	Sequence 50409, A
33	126	58.6	612	5	US-09-791-537-73416	Sequence 73416, A
34	126	58.6	618	5	US-09-791-537-81892	Sequence 81892, A
35	126	58.6	619	5	US-09-791-537-33749	Sequence 33749, A
36	125	58.1	40	5	US-09-791-537-86308	Sequence 86308, A
37	125	58.1	482	5	US-09-791-537-41672	Sequence 41672, A
38	125	58.1	482	5	US-09-791-537-111285	Sequence 111285, A
39	125	58.1	484	5	US-09-791-537-50405	Sequence 50405, A
40	125	58.1	616	5	US-09-791-537-332023	Sequence 332023, A
41	125	58.1	740	5	US-09-791-537-32024	Sequence 32024, A
42	125	58.1	740	5	US-09-791-537-32024	Sequence 32024, A
43	125	58.1	830	5	US-09-791-537-22819	Sequence 22819, A
44	125	58.1	830	5	US-09-791-537-35618	Sequence 35618, A
45	124	57.7	768	5	US-09-791-537-20989	Sequence 20989, A

## ALIGNMENTS

```
RESULT 1
US-09-791-537-152667
; Sequence 152667, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 152667
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667

Query Match      100.0%; Score 215; DB 5; Length 372;
Best Local Similarity 100.0%; Pred. No. 4e-16;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  COPWSCSGHGECEIINNHTCNCVDVGYGPOCOL 34
        |||||||
Db      160 COPWSCSGHGECEIINNHTCNCVDVGYGPOCOL 193

RESULT 2
US-09-791-537-84593
; Sequence 84593, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentln version 3.0
; SEQ ID NO 84593
```

```
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-84593
```

```
Query Match          97.2%; Score 209; DB 5; Length 372;
Best Local Similarity 97.1%; Pred. No. 1.8e-15;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 34
Db      160 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 193
```

```
RESULT      3
US-09-791-537-53485
; Sequence 53485, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53485
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53485
```

```
Query Match          95.3%; Score 205; DB 5; Length 323;
Best Local Similarity 97.0%; Pred. No. 4.5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
Db      173 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 205
```

```
RESULT      4
US-10-211-364-1171
; Sequence 1171, Application US/10211364
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P216CIN
; CURRENT APPLICATION NUMBER: US/10/211,364
; CURRENT FILING DATE: 2002-08-05
; PRIOR APPLICATION NUMBER: 09/760,486
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/218,290
```

```
; PRIOR FILING DATE: 2000-07-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1778
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1171
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-211-364-1171
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
Db      182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 214
```

```
RESULT      5
US-10-212-054-1328
; Sequence 1328, Application US/10212054
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P212CIN
; CURRENT APPLICATION NUMBER: US/10/212,054
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-054-1328
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
Oy      1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33
Db      182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 214
```

```
RESULT      6
US-10-212-778-1158
; Sequence 1158, Application US/10212778
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026CIN
; CURRENT APPLICATION NUMBER: US/10/212,778
; CURRENT FILING DATE: 2002-08-07
; PRIOR APPLICATION NUMBER: 09/758,449
; PRIOR FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: PatentIn Ver. 2.0
```



```
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc-feature
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-212-778-1158
```

```
Query Match          95.3%; Score 205; DB 6; Length 341;
Best Local Similarity 97.0%; Pred. No. 4.7e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 182 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 214
```

```
RESULT 7
US-09-791-537-58446
; Sequence 58446, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58446
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-58446
```

```
Query Match          95.3%; Score 205; DB 5; Length 363;
Best Local Similarity 97.0%; Pred. No. 4.9e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 173 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 205
```

```
RESULT 8
US-09-791-537-22816
; Sequence 22816, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22816
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-22816
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
```

```
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 9
US-09-791-537-42655
; Sequence 42655, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42655
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Macaca mulatta
US-09-791-537-42655
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 10
US-09-791-537-42657
; Sequence 42657, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42657
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-791-537-42657
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 33
|||||
Db 160 COPMSCSGHCEVEIINHTCNCDDVGYGPQCQ 192
```

```
RESULT 11
US-09-791-537-42658
; Sequence 42658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-42658
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Papio hamadryas
US-09-791-537-42658
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 192
```

```

RESULT 12
; Sequence 42659, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42659
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pongo pygmaeus
US-09-791-537-42659
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 192
```

```

RESULT 13
; Sequence 69658, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 69658
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-69658
```

```
Query Match          95.3%; Score 205; DB 5; Length 372;
Best Local Similarity 97.0%; Pred. No. 5e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 33
Db 160 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 192
```

```

RESULT 14
; Sequence 51391, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51391
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-51391
```

```
Query Match          95.3%; Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 33
Db 173 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 205
```

```

RESULT 15
; Sequence 53844, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 53844
; LENGTH: 385
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-53844
```

```
Query Match          95.3%; Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 33
Db 173 CQPMSCGHGECVEIINNHTCNCDDVGYGPOCQ 205
```

```

RESULT 16
US-09-791-537-113060
```

```

: Sequence 113060, Application US//09791537
:
: GENERAL INFORMATION:
:
: APPLICANT: Blomomix, Inc.
:
: APPLICANT: Debe, Derek
:
: APPLICANT: Danzer, Joseph
:
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
:
: TITLE OF INVENTION: METHODS OF USE THEREOF
:
: FILE REFERENCE: 261/210
:
: CURRENT APPLICATION NUMBER: US//09/791,537
:
: CURRENT FILING DATE: 2001-02-22
:
: NUMBER OF SEQ ID NOS: 153055
:
: SOFTWARE: PatentIn version 3.0
:
: SEQ ID NO 113060
:
: LENGTH: 385
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
:
: IS-09-791-537-113060

```

```
Query Match      95.3%   Score 205; DB 5; Length 385;
Best Local Similarity 97.0%; Pred. No. 5.2e-15;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```

RESULT 17
US-09-791-537-50403
: Sequence 50403, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Blonomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 50403
: LENGTH: 370
: TYPE: PRT
: ORGANISM: Bos taurus
: US-09-791-537-50403

```

```

Query Match          88.8%  Score 191:  DB 5:  length 370;
Best Local Similarity 81.8%  Pred. No. 1.8e-13;
Matches 27;  Conservative 0;  Mismatches 6;  Indels 0;  Gaps 0;

QY      1  COPWSCSGHCEVEIINNHTNCDCVGYGPOCQ 33
      |:|||||:|:|:|:|:|:|:|:|:|:|:|:|:|
Db      160  CKPWSCSGHCGCEVINNYTNCDCDLYGYGPCQ 192

RESULT 18
US-09-791-537-132144
: Sequence 132144, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ. ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 132144

```

```

; LENGTH: 376
; TYPE: PRF
; ORGANISM: Oryctolagus cuniculus
US-09-791-537-132144

```

Query Match	84.2%	Score 181	DB 5	Length 376
Best Local Similarity	84.8%	Pred. No. 2.3e-15		
Matches 28; Conservative	3	Mismatches 2	Indels 0	Gaps 0

  

QY	1	CQPMSCSGHGECVELINNHTCNCDVGYGPPCQ	33
hb	160	CHPGSCSGHGECVEYINNHTCSGVGYGPPCQ	192

```

RESULT 19
US-09-791-537-60503
: Sequence 60503, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: patentIn version 3.0
: SEQ ID NO 60503
: LENGTH: 360
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-791-537-60503

```

```

Query Match 78.6%; Score 169; DB 5; Length 360;
Best Local Similarity 81.8%; Pred. NO. 4.7e-11;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 CQWSCSGHGECVELINNHTCNCDVGYGPOCQ 33
   ||| ||| ||| ||| ||| ||| ||| |||
db 160 CQPSGCGNGECVELINNHTCTICDAGYGPCCQ 192

```

```

RESULT 20
US-09-791-537-37750
: Sequence 37750, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Biomomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37750
: LENGTH: 372
: TYPE: PRT
: ORGANISM: Mus musculus
US-09-791-537-37750

```

Query Match	78.6%	Score 169	DB 5	Length 372
Best Local Similarity	81.8%	Pred. No	4.8e-11	
Matches 27; Conservative	1;	Mismatches	5;	Indels 0; Gaps 0;
QY	1	CGPMSCGHGECVETINNHTCNCVDVGYGPPCQ	33	
Db	160	CGPSCNGRGECVETINNHTCNCIDAGYGPCCQ	192	

```
RESULT 21
US-09-791-537-6693
; Sequence 6693, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6693
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Rattus sp
US-09-791-537-6693
```

```
Query Match
Best Local Similarity 74.0%; Score 159; DB 5; Length 372;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 192
```

```
RESULT 22
US-09-791-537-81233
; Sequence 81233, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 81233
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-81233
```

```
Query Match
Best Local Similarity 74.0%; Score 159; DB 5; Length 372;
Matches 26; Conservative 2; Mismatches 5; Indels 0; Gaps 0;
```

```
QY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 160 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 192
```

```
RESULT 23
US-09-791-537-118878
; Sequence 118878, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
```

```
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 118878
; LENGTH: 611
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-118878
```

```
Query Match
Best Local Similarity 60.9%; Score 131; DB 5; Length 611;
Matches 20; Conservative 5; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 144 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 176
```

```
RESULT 24
US-09-791-537-23269
; Sequence 23269, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23269
; LENGTH: 162
; TYPE: PRT
; ORGANISM: pdb 1ESL
US-09-791-537-23269
```

```
Query Match
Best Local Similarity 60.5%; Score 130; DB 5; Length 162;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY 1 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 33
Db 122 CQPMSCGHGECVEITNNHTCNCIDVGYGPQCQ 154
```

```
RESULT 25
PCT-US02-23913-357
; Sequence 357, Application PC/TUS0223913
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc. et al.
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF PROSTATE CANCER
; FILE REFERENCE: MRI-044PC
; CURRENT APPLICATION NUMBER: PCT/US02/23913
; CURRENT FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/307,982
; PRIOR FILING DATE: 2001-07-25
; PRIOR APPLICATION NUMBER: 60/314,356
; PRIOR FILING DATE: 2001-08-22
; PRIOR APPLICATION NUMBER: 60/325,020
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: 60/341,746
; PRIOR FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: 60/362,158
; PRIOR FILING DATE: 2002-03-05
; NUMBER OF SEQ ID NOS: 455
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 357
```

```

; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-23913-357

```

Query Match	60.5%	Score 130;	DB 1;	Length 610;
Best Local Similarity	63.6%	Pred. No. 1.5e-06;		
Matches	21; Conservative	4; Mismatches	8; Indels	0; Gaps

Qy 1 CQPMSCSGHGECVEIINNHTCNCNDVGYYPCQ 33  
| | | | | | | | | | : : : :  
Db 143 CTNTSCSGHGECVETINNYYTCKCDPFGSLKE 175

RESULT 26  
US-09-791-537-22813  
; Sequence 22813, Application US/09791537

: APPLICANT: Danzer, Joseph  
 : TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS  
 : TITLE OF INVENTION: METHODS OF USE THEREOF  
 : FILE REFERENCE: 261/210

```

? CURRENT APPLICATION NUMBER: US/09/791,533
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 153055
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 22813
? LENGTH: 610
? TYPE: PRT
? ORGANISM: Homo sapiens
? OS-09-791-537-22813

```

Query Match	60.5%;	Score 130;	DB 5;	Length 610;
Best Local Similarity	63.6%;	Pred. No. 1.5e-06;		
Matches	21; Conservative	4; Mismatches	8; Indels	0; Gaps 0;

Qy 1 CQPMSCSGHGECVEIINNHTCNCDDVGYGPDQC 33  
| | | | | | | | : : : :  
Db 143 CTNTSCSGHGECEVTINNYTCKCDPGFSLKCE 175

RESULT 27  
US-09-791-537-121834  
; Sequence 121834, Application US/09791537

: APPLICANT: Bionomix, Inc.  
 : APPLICANT: Debe, Derek  
 : APPLICANT: Danzer, Joseph  
 : TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

```

? TITLE OF INVENTION: METHODS OF USE THEREOF
? FILE REFERENCE: 261/210
? CURRENT APPLICATION NUMBER: US/09/791,537
? CURRENT FILING DATE: 2001-02-22
? NUMBER OF SEQ ID NOS: 153055
? SOFTWARE: PatentIn version 3.0
? SEQ ID NO 121834
? LENGTH: 610
? TYPE: PRT
? ORGANISM: Homo sapiens
? SS-09-791-537-121834

```

Query Match	60.5%	Score 130	DB 5	Length 610
Best Local Similarity	63.6%	Pred. No. 1.5e-06		
Matches 21	Conservative	4	Mismatches 8	Indels 0
				Gaps 0

```
QY      1 CQPWSCSGHGECEIINNHTCNCDDVGYGPDCC 33
          | | | | | | | | | | | | : : :
Db     143 CTNTSCSGHGECEVTINNYTCKCDPFGSLKCE 175
```

RESULT 28  
US-10-205-823-357  
: Sequence 357, Application US/10205823

GENERAL INFORMATION:

APPLICANT: Schlagenth, Robert

APPLICANT: Monahan, John E.

APPLICANT: Endege, Wilson O.

APPLICANT: Gannavarapu, Manjula

APPLICANT: Gombarchera, Bella

APPLICANT: Hoersch, Sebastian

APPLICANT: Kamalakar, Shudhangai

APPLICANT: Womey, Angela M.

1 APPLICANT: Galt, Karen  
 2 APPLICANT: Zhao, Xumei  
 3 APPLICANT: Anderson, Dustin  
 4 TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
 5 TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
 6 TITLE OF INVENTION: THERAPY OF PROSTATE CANCER  
 7 FILE REFERENCE: MRI-044  
 8 CURRENT APPLICATION NUMBER: US/10/205,823

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

```

: NUMBER OF SEO ID NOS: 455
: SOFTWARE: Fastseq for Windows Version 4.0.
: SEO ID NO 357
: LENGTH: 610
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-10-205-823-357

```

Query Match	60.5%;	Score 130;	DB 6;	Length 610;
Best Local Similarity	63.68;	Pred. No. 1.5e-06;		
Matches 21; Conservative	4;	Mismatches 8;	Indels 0;	Gaps 0;

Oy 1 COPWSCSGHGEVETINNHTCNCDVGYYGPOCQ 33  
| | | | | | | | | | : :  
Db 143 CTNTSCSGHGEVETINNYTCKDCDPFGSLKCE 175

RESULT 29  
US-09-791-537-50406  
; Sequence 50406, Application US/09791537

```

1 GENERAL INFORMATION:
2 APPLICANT: Bionomix, Inc.
3 APPLICANT: Debe, Derek
4 APPLICANT: Danzer, Joseph
5 TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
6 TITLE OF INVENTION: METHODS OF USE THEREOF
7 FILE REFERENCE: 261/210
8 CURRENT APPLICATION NUMBER: US/09/791,537
9 CURRENT FILING DATE: 2001-02-22
10 NUMBER OF SEQ ID NOS: 153055
11 SOFTWARE: PatentIn version 3.0

```

```

; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-50406

```

Query Match 59.18; Score 127; DB 5; Length 549;









GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:20:37 : Search time 304.64 Seconds  
(without alignments)  
39,284 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193

Perfect score: 215

Sequence: 1 CQPMSCSGHCEVEIINHTCNCDCVGYGQCOL 34

Scoring table:

BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 100%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCrUS.COMB.pep.\*

2: /cgn2\_6/ptodata/2/paa/US06.COMB.pep.\*

3: /cgn2\_6/ptodata/2/paa/US07.COMB.pep.\*

4: /cgn2\_6/ptodata/2/paa/US080.COMB.pep.\*

5: /cgn2\_6/ptodata/2/paa/US081.COMB.pep.\*

6: /cgn2\_6/ptodata/2/paa/US082.COMB.pep.\*

7: /cgn2\_6/ptodata/2/paa/US083.COMB.pep.\*

8: /cgn2\_6/ptodata/2/paa/US084.COMB.pep.\*

9: /cgn2\_6/ptodata/2/paa/US085.COMB.pep.\*

10: /cgn2\_6/ptodata/2/paa/US086.COMB.pep.\*

11: /cgn2\_6/ptodata/2/paa/US087.COMB.pep.\*

12: /cgn2\_6/ptodata/2/paa/US088.COMB.pep.\*

13: /cgn2\_6/ptodata/2/paa/US089.COMB.pep.\*

14: /cgn2\_6/ptodata/2/paa/US090.COMB.pep.\*

15: /cgn2\_6/ptodata/2/paa/US091.COMB.pep.\*

16: /cgn2\_6/ptodata/2/paa/US092.COMB.pep.\*

17: /cgn2\_6/ptodata/2/paa/US093.COMB.pep.\*

18: /cgn2\_6/ptodata/2/paa/US094.COMB.pep.\*

19: /cgn2\_6/ptodata/2/paa/US095.COMB.pep.\*

20: /cgn2\_6/ptodata/2/paa/US096.COMB.pep.\*

21: /cgn2\_6/ptodata/2/paa/US097.COMB.pep.\*

22: /cgn2\_6/ptodata/2/paa/US098.COMB.pep.\*

23: /cgn2\_6/ptodata/2/paa/US099.COMB.pep.\*

24: /cgn2\_6/ptodata/2/paa/US100.COMB.pep.\*

25: /cgn2\_6/ptodata/2/paa/US101.COMB.pep.\*

26: /cgn2\_6/ptodata/2/paa/US60.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	215	100.0	372	15	US-09-119-209-2
2	209	97.2	382	21	US-09-760-475-2123
3	209	97.2	1078	26	US-60-212-659-523
4	205	95.3	341	21	US-09-758-449-1158
5	205	95.3	341	21	US-09-760-443-1328
6	205	95.3	372	1	PCT-US01-26675-3
7	205	95.3	385	1	PCT-US92-03970-2

8	205	95.3	385	1	PCT-US94-00909-2	Sequence 2, Appl1
9	205	95.3 <th>385</th> <th>7</th> <th>US-08-008-459-2</th> <th>Sequence 2, Appl1</th>	385	7	US-08-008-459-2	Sequence 2, Appl1
10	205	95.3 <th>385</th> <th>7</th> <th>US-08-340-539-2</th> <th>Sequence 2, Appl1</th>	385	7	US-08-340-539-2	Sequence 2, Appl1
11	205	95.3 <th>385</th> <th>8</th> <th>US-08-410-569-2</th> <th>Sequence 2, Appl1</th>	385	8	US-08-410-569-2	Sequence 2, Appl1
12	205	95.3 <th>1078</th> <th>26</th> <th>US-60-207-315-428</th> <th>Sequence 428, App</th>	1078	26	US-60-207-315-428	Sequence 428, App
13	205	95.3 <th>1078</th> <th>26</th> <th>US-60-230-435-1751</th> <th>Sequence 1751, App</th>	1078	26	US-60-230-435-1751	Sequence 1751, App
14	169 <th>78.6</th> <th>372</th> <th>15</th> <th>US-09-119-209-4</th> <th>Sequence 4, Appl1</th>	78.6	372	15	US-09-119-209-4	Sequence 4, Appl1
15	130 <th>60.5</th> <th>610</th> <th>1</th> <th>PCT-US99-28965-19</th> <th>Sequence 19, Appl1</th>	60.5	610	1	PCT-US99-28965-19	Sequence 19, Appl1
16	130 <th>60.5</th> <th>610</th> <th>10</th> <th>US-08-657-753-2</th> <th>Sequence 2, Appl1</th>	60.5	610	10	US-08-657-753-2	Sequence 2, Appl1
17	130 <th>60.5</th> <th>610</th> <th>11</th> <th>US-08-770-435-3</th> <th>Sequence 3, Appl1</th>	60.5	610	11	US-08-770-435-3	Sequence 3, Appl1
18	130 <th>60.5</th> <th>610</th> <th>16</th> <th>US-09-266-091-2</th> <th>Sequence 2, Appl1</th>	60.5	610	16	US-09-266-091-2	Sequence 2, Appl1
19	130 <th>60.5</th> <th>610</th> <th>16</th> <th>US-09-266-091A-2</th> <th>Sequence 2, Appl1</th>	60.5	610	16	US-09-266-091A-2	Sequence 2, Appl1
20	130 <th>60.5</th> <th>610</th> <th>21</th> <th>US-09-784-356-122</th> <th>Sequence 122, Appl1</th>	60.5	610	21	US-09-784-356-122	Sequence 122, Appl1
21	130 <th>60.5</th> <th>610</th> <th>22</th> <th>US-09-802-640-36</th> <th>Sequence 36, Appl1</th>	60.5	610	22	US-09-802-640-36	Sequence 36, Appl1
22	130 <th>60.5</th> <th>610</th> <th>22</th> <th>US-09-857-670-19</th> <th>Sequence 19, Appl1</th>	60.5	610	22	US-09-857-670-19	Sequence 19, Appl1
23	130 <th>60.5</th> <th>610</th> <th>24</th> <th>US-10-021-660-122</th> <th>Sequence 122, App</th>	60.5	610	24	US-10-021-660-122	Sequence 122, App
24	125 <th>58.1</th> <th>700</th> <th>26</th> <th>US-60-207-315-467</th> <th>Sequence 467, App</th>	58.1	700	26	US-60-207-315-467	Sequence 467, App
25	125 <th>58.1</th> <th>830</th> <th>1</th> <th>PCT-US94-09395-4</th> <th>Sequence 4, Appl1</th>	58.1	830	1	PCT-US94-09395-4	Sequence 4, Appl1
26	125 <th>58.1</th> <th>830</th> <th>8</th> <th>US-08-449-687B-4</th> <th>Sequence 4, Appl1</th>	58.1	830	8	US-08-449-687B-4	Sequence 4, Appl1
27	125 <th>58.1</th> <th>830</th> <th>24</th> <th>US-10-020-141-10</th> <th>Sequence 10, Appl1</th>	58.1	830	24	US-10-020-141-10	Sequence 10, Appl1
28	120 <th>55.8</th> <th>36</th> <th>26</th> <th>US-60-160-189-8885</th> <th>Sequence 8885, App</th>	55.8	36	26	US-60-160-189-8885	Sequence 8885, App
29	120 <th>55.8</th> <th>37</th> <th>26</th> <th>US-60-160-189-10134</th> <th>Sequence 10134, App</th>	55.8	37	26	US-60-160-189-10134	Sequence 10134, App
30	120 <th>55.8</th> <th>37</th> <th>26</th> <th>US-60-160-203-6335</th> <th>Sequence 6335, App</th>	55.8	37	26	US-60-160-203-6335	Sequence 6335, App
31	120 <th>55.8</th> <th>38</th> <th>26</th> <th>US-60-160-203-5155</th> <th>Sequence 5155, App</th>	55.8	38	26	US-60-160-203-5155	Sequence 5155, App
32	120 <th>55.8</th> <th>38</th> <th>26</th> <th>US-60-169-840-6794</th> <th>Sequence 6794, App</th>	55.8	38	26	US-60-169-840-6794	Sequence 6794, App
33	120 <th>55.8</th> <th>38</th> <th>26</th> <th>US-60-169-867-5901</th> <th>Sequence 5901, App</th>	55.8	38	26	US-60-169-867-5901	Sequence 5901, App
34	98 <th>45.6</th> <th>34</th> <th>26</th> <th>US-60-146-055-529</th> <th>Sequence 529, App</th>	45.6	34	26	US-60-146-055-529	Sequence 529, App
35	98 <th>45.6</th> <th>173</th> <th>26</th> <th>US-60-139-669-665</th> <th>Sequence 665, App</th>	45.6	173	26	US-60-139-669-665	Sequence 665, App
36	98 <th>45.6</th> <th>183</th> <th>26</th> <th>US-60-139-669-596</th> <th>Sequence 596, App</th>	45.6	183	26	US-60-139-669-596	Sequence 596, App
37	98 <th>45.6</th> <th>321</th> <th>16</th> <th>US-09-270-767-33762</th> <th>Sequence 33762, App</th>	45.6	321	16	US-09-270-767-33762	Sequence 33762, App
38	98 <th>45.6</th> <th>321</th> <th>16</th> <th>US-09-270-767-48979</th> <th>Sequence 48979, App</th>	45.6	321	16	US-09-270-767-48979	Sequence 48979, App
39	98 <th>45.6</th> <th>321</th> <th>16</th> <th>US-09-270-8498-194120</th> <th>Sequence 194120, App</th>	45.6	321	16	US-09-270-8498-194120	Sequence 194120, App
40	98 <th>45.6</th> <th>427</th> <th>26</th> <th>US-60-142-886-1034</th> <th>Sequence 1034, App</th>	45.6	427	26	US-60-142-886-1034	Sequence 1034, App
41	98 <th>45.6</th> <th>459</th> <th>26</th> <th>US-60-140-956-1513</th> <th>Sequence 1513, App</th>	45.6	459	26	US-60-140-956-1513	Sequence 1513, App
42	98 <th>45.6</th> <th>496</th> <th>26</th> <th>US-60-145-138-711</th> <th>Sequence 711, App</th>	45.6	496	26	US-60-145-138-711	Sequence 711, App
43	98 <th>45.6</th> <th>3396</th> <th>20</th> <th>US-09-614-150-19575</th> <th>Sequence 19575, App</th>	45.6	3396	20	US-09-614-150-19575	Sequence 19575, App
44	98 <th>45.6</th> <th>3396</th> <th>26</th> <th>US-60-167-217-19718</th> <th>Sequence 19718, App</th>	45.6	3396	26	US-60-167-217-19718	Sequence 19718, App
45	98 <th>45.6</th> <th>3396</th> <th>26</th> <th>US-60-191-637-19634</th> <th>Sequence 19634, App</th>	45.6	3396	26	US-60-191-637-19634	Sequence 19634, App

#### ALIGNMENTS

RESULT 1

US-09-119-209-2

Sequence 2, Application US/09119209

GENERAL INFORMATION:

APPLICANT: LASKY, LAURENCE A.

APPLICANT: SPACHELL, SCOTT E.

APPLICANT: ROSEN, STEVEN D.

APPLICANT: SINGER, MARK S.

APPLICANT: YEDNOCK, TED A.

TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Winpatin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/119,209

FILING DATE: 20-Jul-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/513278

FILING DATE: 10-AUG-1995

```
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: 08/059027
;; FILING DATE: 6-MAY-1993
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: 07/786149
;; FILING DATE: 31-OCT-1991
;; PRIORITY APPLICATION DATA:
;; APPLICATION NUMBER: 07/315015
;; FILING DATE: 23-FEB-1989
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Love, Richard B.
;; REGISTRATION NUMBER: 34,659
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 650/225-5530
;; TELEFAX: 650/952-9881
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 372 amino acids
;; TYPE: Amino Acid
;; TOPOLOGY: Linear
US-09-119-209-2
```

```
Query Match 100.0%; Score 215; DB 15; Length 372;
Best Local Similarity 100.0%; Pred. No. 4.5e-17;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 34
Db 160 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 193
```

```
RESULT 2
US-09-760-475-2123
; Sequence 2123, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 2123
; LENGTH: 382
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-760-475-2123
```

```
Query Match 97.2%; Score 209; DB 21; Length 382;
Best Local Similarity 97.1%; Pred. No. 2.4e-16;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 34
Db 170 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 203
```

```
RESULT 3
US-60-212-659-523
; Sequence 523, Application US/60212659
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CLO00674
; CURRENT APPLICATION NUMBER: US/60/212,659
; CURRENT FILING DATE: 2000-06-19
; NUMBER OF SEQ ID NOS: 879
```

```
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 523
;; LENGTH: 1078
;; TYPE: PRT
;; ORGANISM: HUMAN
US-60-212-659-523
```

```
Query Match 97.2%; Score 209; DB 26; Length 1078;
Best Local Similarity 97.1%; Pred. No. 6e-16;
Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 34
Db 742 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 775
```

```
RESULT 4
US-09-758-449-1158
; Sequence 1158, Application US/09758449
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM026
; CURRENT APPLICATION NUMBER: US/09/758,449
; CURRENT FILING DATE: 2001-01-11
; PRIORITY APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 1478
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1158
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-449-1158
```

```
Query Match 95.3%; Score 205; DB 21; Length 341;
Best Local Similarity 97.0%; Pred. No. 6.3e-16;
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 33
Db 182 CQPMSCSGHGECEIINNHTCNCDDVGYGPQCOL 214
```

```
RESULT 5
US-09-760-443-1328
; Sequence 1328, Application US/09760443
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P7212
; CURRENT APPLICATION NUMBER: US/09/760,443
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2164
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1328
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (215)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-443-1328
```

Query Match 95.3%; Score 205; DB 21; Length 341;  
Best Local Similarity 97.0%; Pred. No. 6.3e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33  
Db 182 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 214

RESULT 6  
PCT-US01-26675-3  
Sequence 3, Application PC/TUS0126675  
GENERAL INFORMATION:  
APPLICANT: Genaisance Pharmaceuticals, Inc.  
APPLICANT: Anastasio, Allison E  
APPLICANT: Biegleski, Karyn M  
APPLICANT: Kishy, Stefanie E  
APPLICANT: Kumar, Anant Madan  
TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE  
FILE REFERENCE: SELL MWH1116-PCF  
CURRENT APPLICATION NUMBER: PCT/US01/26675  
CURRENT FILING DATE: 2001-08-27  
PRIOR APPLICATION NUMBER: 60/228,262  
PRIOR FILING DATE: 2000-08-25  
NUMBER OF SEQ ID NOS: 101  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3  
LENGTH: 372  
TYPE: PRT  
ORGANISM: Homo sapiens  
PCT-US01-26675-3

Query Match 95.3%; Score 205; DB 1; Length 372;  
Best Local Similarity 97.0%; Pred. No. 6.9e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33  
Db 160 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 192

RESULT 7  
PCT-US92-03970-2  
Sequence 2, Application PC/TUS9203970  
GENERAL INFORMATION:  
APPLICANT: Dana-Farber Cancer Institute, Inc.  
TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE  
TITLE OF INVENTION: PROTEIN  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/03970  
FILING DATE: 19920513  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Helne, Holliday C.  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: DFCI-152Bq9

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
TELEX: 940675  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-03970-2

Query Match 95.3%; Score 205; DB 1; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7.1e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 33  
Db 173 CQPMSCSGHGECEIINNHTCNCDDVGYGPOCQ 205

RESULT 8  
PCT-US94-00909-2  
Sequence 2, Application PC/TUS9400909  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING  
TITLE OF INVENTION: AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/00909  
APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/983,606  
FILING DATE: 30-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/962,483  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,608  
FILING DATE: 03-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/737,092  
FILING DATE: 29-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/730,503  
FILING DATE: 08-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/700,773  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/313,109  
FILING DATE: 21-FEB-1989  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-00909-2

Query Match 95.3%; Score 205; DB 1; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7.1e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPMSCGHECEVEIINNHTCNCMDVGYGPOCQ 33  
Db 173 COPMSCGHECEVEIINNHTCNCMDVGYGPOCQ 205

## RESULT 9

US-08-008-459-2  
; Sequence 2, Application US/08008459  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/008,459  
; FILING DATE: 25-JAN-1993  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/983,606  
; FILING DATE: 30-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/862,483  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,608  
; FILING DATE: 03-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/737,092  
; FILING DATE: 29-JUL-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/730,503  
; FILING DATE: 08-JUL-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/700,773  
; FILING DATE: 15-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/313,109  
; FILING DATE: 21-FEB-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heine, Holliday C.  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: DFCI-318XX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; TELETYPE: 940675  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-008-459-2

Query Match 95.3%; Score 205; DB 4; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7.le-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPMSCGHECEVEIINNHTCNCMDVGYGPOCQ 33  
Db 173 COPMSCGHECEVEIINNHTCNCMDVGYGPOCQ 205

## RESULT 10

US-08-340-539-2  
; Sequence 2, Application US/08340539  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,539  
; FILING DATE: 16-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,459  
; FILING DATE: 25-JAN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/983,606  
; FILING DATE: 30-NOV-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/862,483  
; FILING DATE: 02-APR-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,608  
; FILING DATE: 03-OCT-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/737,092  
; FILING DATE: 29-JUL-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/730,503  
; FILING DATE: 08-JUL-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/700,773  
; FILING DATE: 15-MAY-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/313,109  
; FILING DATE: 21-FEB-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heine, Holliday C.  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: DFCI-318XX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; TELETYPE: 940675  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-340-539-2

Query Match 95.3%; Score 205; DB 7; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7.le-16;

Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 33  
DB 173 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 205

RESULT 11  
US-08-410-569-2

; Sequence 2, Application US/08410569  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Sperdin, Olivier G.  
; TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)  
; TITLE OF INVENTION: AND LIGAND THEREOF  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/08/410,569  
; APPLICATION NUMBER: US/08/410,569  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/770,608  
; FILING DATE: 03-OCT-1991  
; APPLICATION NUMBER: US 07/700,773  
; FILING DATE: 15-MAY-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Helne, Holliday C.  
; REGISTRATION NUMBER: 34,346  
; REFERENCE/DOCKET NUMBER: DFCG-152EX  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-2290  
; TELEFAX: (617) 451-0313  
; TELEX: 940675  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-410-569-2

Query Match 95.3%; Score 205; DB 8; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7.1e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 33  
DB 173 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 205

RESULT 12  
US-60-207-315-428

; Sequence 428, Application US/60207315  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: CLO00601  
; CURRENT APPLICATION NUMBER: US/60/207,315

; CURRENT FILING DATE: 2000-05-30

; NUMBER OF SEQ ID NOS: 528

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 428

; LENGTH: 1078

; TYPE: PRT

; ORGANISM: HUMAN

US-60-207-315-428

Query Match 95.3%; Score 205; DB 26; Length 1078;  
Best Local Similarity 97.0%; Pred. No. 1.8e-15;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 33  
DB 742 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 774

RESULT 13  
US-60-230-435-1751

; Sequence 1751, Application US/60230435  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: CLO00768  
; CURRENT APPLICATION NUMBER: US/60/230,435  
; CURRENT FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 2991  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1751  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-230-435-1751

Query Match 95.3%; Score 205; DB 26; Length 1078;  
Best Local Similarity 97.0%; Pred. No. 1.8e-15;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 33  
DB 742 COPWSCSGHCEVEIINNHTCNCVDGYGPGCQ 774

RESULT 14  
US-09-119-209-4

; Sequence 4, Application US/09119209  
; GENERAL INFORMATION:  
; APPLICANT: LASKY, LAURENCE A.  
; APPLICANT: SPACHELL, SCOTT E.  
; APPLICANT: ROSEN, STEVEN D.  
; APPLICANT: SINGER, MARK S.  
; APPLICANT: YEDNOCK, TED A.  
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Genentech, Inc.  
; STREET: 1 DNA Way  
; CITY: South San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94080  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: WinPatIn (Genentech)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/119,209

```

; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
US-09-119-209-4
```

```
Query Match          78.6%; Score 169; DB 15; Length 372;
Best Local Similarity 81.8%; Pred. No. 1.3e-11;
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 33
Db      160 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 192
```

```

RESULT 15
PCT-US99-28965-19
; Sequence 19, Application PC/TUS9928965
; GENERAL INFORMATION:
; APPLICANT: Monla, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0424
; CURRENT FILING DATE: 1999-12-08
; EARLIER FILING DATE: 1998-12-10
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US99-28965-19
```

```
Query Match          60.5%; Score 130; DB 1; Length 610;
Best Local Similarity 63.6%; Pred. No. 8.6e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 33
Db      143 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 175
```

```

RESULT 16
US-08-657-753-2
; Sequence 2, Application US/08657753
```

```

; GENERAL INFORMATION:
; APPLICANT: Klimuk, Sandra K.
; APPLICANT: Semple, Sean C.
; APPLICANT: Scherter, Peter
; APPLICANT: Hope, Michael J.
; TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESS: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/657,753
; FILING DATE: Not yet assigned
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kezer, William B.
; REGISTRATION NUMBER: 37,369
; REFERENCE/DOCKET NUMBER: 16303-003600
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-657-753-2
```

```
Query Match          60.5%; Score 130; DB 10; Length 610;
Best Local Similarity 63.6%; Pred. No. 8.6e-07;
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 33
Db      143 CQPMSCSGHGECEIINNHTCNCVGYGPOCQ 175
```

```

RESULT 17
US-08-770-435-3.
; Sequence 3, Application US/08770435
; GENERAL INFORMATION:
; APPLICANT: Gimbrone, Jr., Michael A.
; APPLICANT: Bevilacqua, Michael P.
; TITLE OF INVENTION: Antibodies Specific For E-selectin And The Uses
; TITLE OF INVENTION: Thereof
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESS: STERN, KESSLER, GOLDSTEIN & FOX
; STREET: 1100 New York Ave., NW
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,435
; FILING DATE: herewith
; CLASSIFICATION: 424
```



Query Match	60.5%	Score 130;	DB 21;	Length 610;
Best Local Similarity	63.6%	Pred. No. 8.6e-07;		
Matches 21; Conservative	4;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 CQPMSCSGHGCEVEIINNHTCNCDDVGYXPQC 33  
| | | | | | | | | | | | | | | | | | : | : | :  
Db 143 CTNTSCSGHGCEVETINNYTCKDCDPFSGLKE 175

RESULT 21  
US-09-802-640-36

```

? Sequence 36, Application US/09802640
? GENERAL INFORMATION:
?
? APPLICANT: Braun, Andreas
? APPLICANT: Bonsal Aruna
? APPLICANT: Kleyn Patrick
? TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
? TITLE OF INVENTION: CARDIOVASCULAR DISEASE AND THEIR USE
? FILE REFERENCE: 24736-2048
? CURRENT APPLICATION NUMBER: US/09/802,640
? CURRENT FILING DATE: 2001-03-09
?
? NUMBER OF SEQ ID NOS: 122
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 36
? LENGTH: 610
? TYPE: PRT
? ORGANISM: Homo sapien
?
? US-09-802-640-36

```

Query Match	60.5%;	Score 130;	DB 22;	Length 610;
Best Local Similarity	63.6%;	Pred. No. 8.6e-07;		
Matches 21; Conservative	4;	Mismatches 8;	Indels 0;	Gaps 0

QY 1 CQPMSCSGHGEVETINNHTCNDVGYGPQCQ 33  
1 | | | | | | | | | | | | | | | | : | : | :  
Db 143 CTNTSCSGHGEVETINNHTCKDPPGSLKCE 175

```

RESULT 22
US-09-857-670-19
Sequence 19, Application US/09857670
GENERAL INFORMATION:
APPLICANT: Monia, Brett P.
APPLICANT: Xu, Xiaoxing S.
APPLICANT: Isis Pharmaceuticals, Inc.
TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
FILE REFERENCE: ISPH-0424
CURRENT APPLICATION NUMBER: US/09/857,670
CURRENT FILING DATE: 2001-09-10
PRIOR APPLICATION NUMBER: US 09/209,668
PRIOR FILING DATE: 1998-12-10
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 19
LENGTH: 610
TYPE: PRN
ORGANISM: Homo sapiens
US-09-857-670-19

```

Query Match	60.5%	Score 130;	DB 22;	Length 610;
Best local Similarity	63.6%	Pred. No. 8.6e-07;		
Matches 21; Conservative	4;	Mismatches 8;	Indels 0;	Gaps 0;

OY 1 CQPMSCSGHGECVEIINNHTCNCDDVGGYGPCQ 33  
| | | | | | | | : : : : : :  
Db 143 CTNTSCSGHGECVETINNYTCKDCDPFSGLKCE 175

RESULT 23  
US-10-021-660-122

```

; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: For Director

```

```

? TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis',
? Technology, Inc.
? TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis
? TITLE OF INVENTION: Modulators
? FILE REFERENCE: 018501-00071005
? CURRENT APPLICATION NUMBER: US/10/021,660
? CURRENT FILING DATE: 2001-12-06
? PRIOR APPLICATION NUMBER: US/09/7784,356
? PRIOR FILING DATE: 2001-02-14
? PRIOR APPLICATION NUMBER: US 09/637,977
? PRIOR FILING DATE: 2000-08-11
? NUMBER OF SEQ ID NOS: 135
? SOFTWARE: FastSeq for Windows Version 3.0
? SEQ ID NO 122
? LENGTH: 610
? TYPE: prt
? ORGANISM: Homo sapiens
? US-10-021,660-122

```

Query Match	60.58;	Score 130;	DB 24;	Length 610;
Best Local Similarity	63.68;	Pred. No. 8.6e-07;		
Matches 21; Conservative	4;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 CQPMSCSGHGEVEIINNHTCMCDVGYYPCQC 33  
| | | | | : | | : | :  
Db 143 CTNTSCSGHGEVETINNYTCKDCDPFSLKCE 175

```

RESULT 24
US-60-207-315-467
: Sequence 467, Application US/60207315
: GENERAL INFORMATION:
: APPLICANT: Beasley, Ellen
: TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,
: TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
: TITLE OF INVENTION: USES THEREOF
: FILE REFERENCE: CLO00601
: CURRENT APPLICATION NUMBER: US/60/207,315
: CURRENT FILING DATE: 2000-05-30
: NUMBER OF SEQ ID NOS: 528
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 467
: LENGTH: 700
: TYPE: PRT
: ORGANISM: HUMAN
: FEATURE:
: NAME/KEY: VARIANT
: LOCATION: (1)..(700)
: OTHER INFORMATION: Xaa = Any Amino Acid
: US-60-207-315-467

```

Query Match	58.1%;	Score 125;	DB 26;	Length 700;
Best Local Similarity	57.6%;	Pred. No. 3.8e-06;		
Matches 19; Conservative	6;	Mismatches 8;	Indels 0;	Gaps 0;

QY 1 CQPMSCSGHGEVETINNHTCNCDVGYYGPQC 33  
11 111 111:1 1:11:1 1:111:1  
Db 155 CQDMSCSKQGECLTIGNYTCSCYPFGYPECE 187

```

RESULT      25
PCT-US94-09395-4
; Sequence 4, Application PC/TUS9409395
; GENERAL INFORMATION:

```



APPLICANT: Board of Regents of the University of Oklahoma  
TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 1100 Peachtree Street, Suite 2800  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-4530  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/09395  
FILING DATE: 19-AUG-1994  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)-815-6508  
TELEFAX: (404)-815-6555  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 830 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
PCT-US94-09395-4

Query Match 58.1%; Score 125; DB 1; Length 830;  
Best Local Similarity 57.6%; Pred. No. 4.5e-06;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQCQ 33  
Db 163 CODMSCSKOGCLETIGNYTCSCYPFGPECE 195

RESULT 26  
US-08-449-687B-4  
Sequence 4, Application US/08449687B  
GENERAL INFORMATION:  
APPLICANT: McEver, Rodger P.  
APPLICANT: Pan, Junliang  
TITLE OF INVENTION: Expression Control Sequences of the  
NUMBER OF SEQUENCES: 17  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patrea L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street  
CITY: Atlanta  
STATE: GA  
COUNTRY: USA  
ZIP: 30309-3450  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/449,687B  
FILING DATE: 24-MAY-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/110,158  
FILING DATE: 20-AUG-1993

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/320,408  
FILING DATE: 08-MAR-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (404)873-8794  
TELEFAX: (404)873-8795  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 830 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-449-687B-4

Query Match 58.1%; Score 125; DB 8; Length 830;  
Best Local Similarity 57.6%; Pred. No. 4.5e-06;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQCQ 33  
Db 163 CODMSCSKOGCLETIGNYTCSCYPFGPECE 195

RESULT 27  
US-10-020-141-10  
Sequence 10, Application US/10020141  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Jeanette  
APPLICANT: Ableson, Allen  
TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE  
FILE REFERENCE: MMJ-002  
CURRENT APPLICATION NUMBER: US/10/020,141  
CURRENT FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/313,097  
PRIOR FILING DATE: 2001-08-16  
PRIOR APPLICATION NUMBER: US 60/327,485  
PRIOR FILING DATE: 2001-10-05  
NUMBER OF SEQ ID NOS: 21  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 830  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-020-141-10

Query Match 58.1%; Score 125; DB 24; Length 830;  
Best Local Similarity 57.6%; Pred. No. 4.5e-06;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Qy 1 COPWSCSGHCEVEIINHTCNCDDVGYGPQCQ 33  
Db 163 CODMSCSKOGCLETIGNYTCSCYPFGPECE 195

RESULT 28  
US-60-160-189-8885  
Sequence 8885, Application US/60160189  
GENERAL INFORMATION:  
APPLICANT: BONA221, VIVIEN  
TITLE OF INVENTION: ISOLATED HUMAN DRUG TARGET PROTEINS,  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN DRUG TARGET PROTEINS  
FILE REFERENCE: CLO00112  
CURRENT APPLICATION NUMBER: US/60/160,189  
CURRENT FILING DATE: 1999-10-19  
NUMBER OF SEQ ID NOS: 10162  
SOFTWARE: FastSeq for Windows Version 4.0





```
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 48979
; LENGTH: 321
; TYPE: PRF
; ORGANISM: Drosophila melanogaster
US-09-270-767-48979
```

```
Query Match          45.6%; Score 98; DB 16; Length 321;
Best Local Similarity 39.4%; Pred. No. 0.003;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1:||||:||||:|1|1|
Db      138 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 170
```

```
RESULT 39
US-09-270-849B-194120
; Sequence 194120, Application US/09270849B
; GENERAL INFORMATION:
; APPLICANT: Swimmer et al.
; TITLE OF INVENTION: Insect genome survey devices
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/270,849B
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 195450
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 194120
; LENGTH: 321
; TYPE: PRF
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-270-849B-194120
```

```
Query Match          45.6%; Score 98; DB 16; Length 321;
Best Local Similarity 39.4%; Pred. No. 0.003;
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
```

```
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1:||||:||||:|1|1|
Db      138 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 170
```

```
RESULT 40
US-60-142-896-1034
; Sequence 1034, Application US/60142896
; GENERAL INFORMATION:
; APPLICANT: Kerlaavage, Anthony
; TITLE OF INVENTION: ISOLATED SECRETED PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING SECRETED PROTEINS, AND USES THEREOF
; FILE REFERENCE: CL000052
; CURRENT APPLICATION NUMBER: US/60/142,896
; CURRENT FILING DATE: 1999-07-08
; NUMBER OF SEQ ID NOS: 1584
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1034
; LENGTH: 427
; TYPE: PRF
; ORGANISM: Drosophila
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(427)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-142-896-1034
```

```
Query Match          45.6%; Score 98; DB 26; Length 427;
Best Local Similarity 39.4%; Pred. No. 0.0039;
```

```
Matches 13; Conservative 9; Mismatches 11; Indels 0; Gaps 0;
QY      1 CQPMSCSGHGEVCEIINNHTCNCVDVGYGPQCQ 33
      1: 1:||||:||||:|1|1|
Db      72 CENRPCANEGTCVDLVGYSCNCEPGYTGKNCQ 104
```

Search completed: September 7, 2002, 10:20:38  
Job time: 486 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:42 ; Search time 28.63 Seconds  
(without alignments)  
29.007 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_160\_193

Perfect score: 215

Sequence: 1 CQWCSGHCVEIINNHTCNCDCVGYGQCQL 34

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	215	100.0	372	2	US-08-513-278-2
2	215	100.0	372	1	Sequence 2, Appl
3	205	95.3	385	1	Patent No. 5514582
4	205	95.3	385	2	Sequence 2, Appl
5	191	88.8	36	1	US-08-461-592B-2
6	169	78.6	34	6	Sequence 22, Appl
7	169	78.6	36	1	Patent No. 5514582
8	169	78.6	372	2	Sequence 4, Appl
9	169	78.6	372	6	Sequence 20, Appl
10	159	74.0	36	1	Patent No. 5514582
11	130	60.5	36	1	Sequence 21, Appl
12	130	60.5	610	1	Sequence 23, Appl
13	130	60.5	610	3	Sequence 3, Appl
14	130	60.5	610	1	Sequence 19, Appl
15	130	60.5	610	4	Sequence 89, Appl
16	127	59.1	610	6	Sequence 9, Appl
17	126	58.6	36	1	Patent No. 5217870
18	125	58.1	36	1	Sequence 24, Appl
19	125	58.1	36	1	Sequence 26, Appl
20	125	58.1	484	2	Sequence 2, Appl
21	125	58.1	484	3	Sequence 9, Appl
22	125	58.1	830	1	Sequence 4, Appl
23	125	58.1	830	5	Sequence 2, Appl
24	123	57.2	830	6	Patent No. 5378464
25	120	55.8	36	1	Sequence 25, Appl
26	93	43.3	36	1	Sequence 27, Appl
27	93	43.3	642	3	Sequence 3, Appl

28	93	43.3	702	4	US-09-068-740A-4
29	93	43.3	723	1	US-09-068-740A-9
30	93	43.3	2523	4	US-08-185-432-18
31	90	41.9	721	3	US-08-872-855-7
32	90	41.9	721	3	US-08-872-855-7
33	89	41.4	2471	1	US-08-981-392-5
34	89	41.4	2471	1	US-08-185-432-16
35	89	41.4	2471	1	US-08-083-590A-19
36	88	40.9	2471	3	US-08-533-364-19
37	87	40.5	713	3	US-08-872-855-5
38	87	40.5	728	4	US-08-981-392-2
39	86	40.0	34	6	US-08-872-855-8
40	86	40.0	2703	1	US-08-185-432-19
41	84	39.1	717	3	US-08-872-855-9
42	84	39.1	1523	4	US-09-182-024A-2
43	84	39.1	2199	5	PCT-US95-11684-2
44	83	38.6	1139	1	US-08-537-210A-4
45	83	38.6	1139	4	US-09-113-825-4

#### ALIGNMENTS

RESULT 1  
US-08-513-278-2  
Sequence 2, Application US/08513278  
Patent No. 5840844  
GENERAL INFORMATION:  
APPLICANT: LASKY, LAURENCE A.  
APPLICANT: STACHELL, SCOTT E.  
APPLICANT: ROSEN, STEVEN D.  
APPLICANT: SINGER, MARK S.  
APPLICANT: YEDNICK, TED A.  
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/513,278  
FILING DATE: 10-AUG-1995  
CLASSIFICATION: 5530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/059027  
FILING DATE: 06-MAY-1993  
APPLICATION NUMBER: 07/786149  
FILING DATE: 31-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/315015  
FILING DATE: 23-FEB-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 565DICI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-513-278-2

Sequence 4, Appl  
Sequence 9, Appl  
Sequence 18, Appl  
Sequence 7, Appl  
Sequence 5, Appl  
Sequence 16, Appl  
Sequence 19, Appl  
Sequence 5, Appl  
Sequence 2, Appl  
Sequence 8, Appl  
Patent No. 5514582  
Sequence 19, Appl  
Sequence 2, Appl  
Sequence 9, Appl  
Sequence 2, Appl  
Sequence 4, Appl  
Sequence 4, Appl

Query Match 100.0%; Score 215; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 5,6e-17;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 34  
|||||  
DB 160 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 193

RESULT 2  
5514582-2  
; Patent No. 5514582  
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
; IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 43  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/185,670  
; FILING DATE: 21-JAN-1994  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 986,931  
; FILING DATE: 08-DEC-1992  
; APPLICATION NUMBER: 808,122  
; FILING DATE: 16-DEC-1991  
; APPLICATION NUMBER: 440,625  
; FILING DATE: 22-NOV-1989  
; APPLICATION NUMBER: 315,015  
; FILING DATE: 23-FEB-1989  
; SEQ ID NO: 2:  
; LENGTH: 372  
5514582-2

Query Match 100.0%; Score 215; DB 6; Length 372;  
Best Local Similarity 100.0%; Pred. No. 5,6e-17;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 34  
|||||  
DB 160 COPWSCSGHGEVEIINNHTCNCNVGYGPOCOL 193

RESULT 3  
US-08-340-539A-2  
; Sequence 2, Application US/08340539A  
; Patent No. 5808025  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,539A  
; FILING DATE: 16-NOV-1994  
; CLASSIFICATION: 514  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,459  
; FILING DATE: 25-JAN-1993

ATTORNEY/AGENT INFORMATION:  
; NAME: Gunnison, Jane  
; REGISTRATION NUMBER: 38,479  
; REFERENCE/DOCKET NUMBER: CG-104 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-340-539A-2

Query Match 95.3%; Score 205; DB 1; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7,6e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVEIINNHTCNCNVGYGPOCQ 33  
|||||  
DB 173 COPWSCSGHGEVEIINNHTCNCNVGYGPOCQ 205

RESULT 4  
US-08-461-592B-2  
; Sequence 2, Application US/08461592B  
; Patent No. 5834425  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
; NUMBER OF SEQUENCES: 11  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Weingarten, Schurigin, Gagnebin & Hayes  
; STREET: Ten Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,592B  
; FILING DATE:  
; CLASSIFICATION: 514  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/340,539  
; FILING DATE: 16-NOV-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,459  
; FILING DATE: 25-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: James F. Haley, Jr.  
; REGISTRATION NUMBER: 27,794  
; REFERENCE/DOCKET NUMBER: CG-104  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 596-9000  
; TELEFAX: (212) 596-9090  
; TELETYPE: 14-8367  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 385 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-592B-2

Query Match 95.3%; Score 205; DB 2; Length 385;  
Best Local Similarity 97.0%; Pred. No. 7, 6e-16;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPWSCSGHGECEVEIINHTCNCDDVGYGPQO 33  
|||||  
Db 173 COPWSCSGHGECEVEIINHTCNCDDVGYGPQO 205

## RESULT 5

US-08-340-539A-22  
; Sequence 22, Application US/08340539A  
; Patent No. 5808025

## GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340, 539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-22

Query Match 88.8%; Score 191; DB 1; Length 36;  
Best Local Similarity 81.8%; Pred. No. 2, 7e-15;  
Matches 27; Conservative 6; Mismatches 0; Indels 0; Gaps 0;

Qy 1 COPWSCSGHGECEVEIINHTCNCDDVGYGPQO 33  
|:|||||:|||||:|||||:|||||:|||||:|  
Db 3 CKPWSCSGHGCEVEIINHTCNCDDVGYGPQO 35

## RESULT 6

5514582-21  
; Patent No. 5514582  
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
; IMMUNOGLOBULINS  
; NUMBER OF SEQUENCES: 43  
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,670  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 986,931  
FILING DATE: 08-DEC-1992  
APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989  
SEQ ID NO: 21:  
LENGTH: 34  
5514582-21

Query Match 78.6%; Score 169; DB 6; Length 34;  
Best Local Similarity 81.8%; Pred. No. 7, 5e-13;  
Matches 27; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 COPWSCSGHGECEVEIINHTCNCDDVGYGPQO 33  
|||:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|:|  
Db 1 COPWSCSGHGECEVEIINHTCNCDDVGYGPQO 33

## RESULT 7

US-08-340-539A-20  
; Sequence 20, Application US/08340539A  
; Patent No. 5808025

## GENERAL INFORMATION:

APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340, 539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 20:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-20

Query Match 78.6%; Score 169; DB 1; Length 36;  
Best Local Similarity 81.8%; Pred. No. 7, 9e-13;





Query Match	74.0%	Score 159;	DB 1;	Length 36;
Best Local Similarity	78.8%	Pred. No. 1e-11;		
Matches	26;	Conservative	2;	Mismatches 5;
				Indels 0;
				Gaps 0;

```

QY      1 CQPMSCSGHGECVEETINNHTCNCDDVGYGPGCQ 33
      ||| ||| : ||| ||| ||| |||
DB      3 CQPESCNRHGECVEETINNHTCICDPGYGPGCQ 35

```

```

US-08-340-539A-23
: RESULT 11
: Sequence 23, Application US/08340539A
: Patent No. 5808025
:
: GENERAL INFORMATION:
: APPLICANT: Tedder, Thomas F.
: APPLICANT: Kansas, Geoffrey S.
: TITLE OF INVENTION: CHEMICAL SELECTIONS AS SIMULTANEOUS
: TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
: NUMBER OF SEQUENCES: 28
: CORRESPONDENCE ADDRESS:
: ADDRESSER: FISH & NEAVE
: STREET: 1251 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10020
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/340,539A
: FILING DATE: 16-NOV-1994
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/008,459
: FILING DATE: 25-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Gunnison, Jane
: REGISTRATION NUMBER: 38,479
: REFERENCE/DOCKET NUMBER: CG-104 CON
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-596-9090
:
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 36 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-340-539A-23

```

Query March 1:	60.5%:	Score 130:	DB 1:	Length 36:
Best Local Similarity	63.6%:	Pred. No. 1.8e-08:		
Matches 21:	Conservative 4:	Mismatches 8:	Indels 0:	Gaps 0:
OY	1	CQPMSCSGHGECVELINNHTCNCVDYGYGPOCQ	33	
cb	3	CTNTSCSGHGECVELINNHTCTCKDPEGFGLACE	35	

RESULT 12  
 US-08-365-470-3  
 Sequence No. 5635991  
 Patent No. 5635991  
 GENERAL INFORMATION:  
 APPLICANT: Glubrone, Jr., Michael A.  
 TITLE OF INVENTION: Antibodies Specific For E-selectin And The Uses  
 TITLE OF INVENTION: Thereof

NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX  
STREET: 1100 NEW YORK AVE., NW  
CITY: Washington

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/365,470
FILING DATE: herewith
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/102,510
FILING DATE: 05-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/850,802
FILING DATE: 13-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Markowicz, Karen R.
REGISTRATION NUMBER: 36,351
REFERENCE/DOCKET NUMBER: 0627.1350003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2540
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
Type: amino acid
TOPOLOGY: linear
US-08-365-470-3

```

Query Match	60.5%	Score 130;	DB 1;	Length 610;
Best Local	63.6%	Pred. No. 3e-07;		
Matches 21;	Conservative 4;	Mismatches 8;	Indels 0;	Gaps 0;

```

RESULT 13
US-09-209-668-19
: Sequence 19, Application US/09209668A
: Patent No. 6114517
: GENERAL INFORMATION:
: APPLICANT: Monla, Brett P.
: APPLICANT: Xu, Xiaoxing S.
: TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
: TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
: FILE REFERENCE: ISPH-0336
: CURRENT APPLICATION NUMBER: US/09/209,668A
: CURRENT FILING DATE: 1998-12-10
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 19
: LENGTH: 610
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-09-209-668-19

```

Query Match	60.5%	Score 130	DB 3	Length 610
Best Local Similarity	63.6%	Pred. No. 3e-07		
Matches 21	Conservative 4	Mismatches 8	Indels 0	Gaps 0
1	CQPMSCSGHGCEVLEIINNHTCNCDDVGYGPOCQ	33		

Db 143 CTNTSCGHEGCEVTINNTYCKDCDPGFSGLKCE 175

RESULT 14  
US-09-009-490A-89  
; Sequence 89, Application US/09009490A  
; Patent No. 6300491  
; GENERAL INFORMATION:  
; APPLICANT: Bennett and Mirabelli  
; TITLE OF INVENTION: Oligonucleotide Modulation  
; TITLE OF INVENTION: of Cell Adhesion  
; NUMBER OF SEQUENCES: 95  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Law Office of Jane Massey Licata  
; STREET: 66 East Main Street  
; CITY: Marlton  
; STATE: NJ  
; COUNTRY: USA  
; ZIP: 08053  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: DISKETTE, 3.5 INCH, 1.44 MB STORAGE  
; COMPUTER: IBM PS/2  
; OPERATING SYSTEM: Windows 95  
; SOFTWARE: WORDPERFECT 6.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/009,490A  
; FILING DATE: January 20, 1998  
; CLASSIFICATION: 514  
; APPLICATION NUMBER: 440,740  
; FILING DATE: May 12, 1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 063,167  
; FILING DATE: May 17, 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 969,151  
; FILING DATE: February 10, 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 007,997  
; FILING DATE: January 20, 1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 939,855  
; FILING DATE: September 2, 1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 567,286  
; FILING DATE: August 14, 1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Jane Massey Licata  
; REGISTRATION NUMBER: 32,257  
; REFERENCE/DOCKET NUMBER: ISPH-0268  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (609) 810-1515  
; TELEFAX: (609) 810-1454  
; INFORMATION FOR SEQ ID NO: 89:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 610  
; TYPE: Amino Acid  
; STRANDEDNESS: Single  
; TOPOLOGY: Linear  
; ANTI-SENSE: no  
US-09-009-490A-89

Query Match 60.5%; Score 130; DB 4; Length 610;  
Best Local Similarity 63.6%; Pred. No. 3e-07;  
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWCSGHEGCEVTINNTYCKDCDPGFSGLKCE 33  
Db 143 CTNTSCGHEGCEVTINNTYCKDCDPGFSGLKCE 175

RESULT 15  
5217870-2  
; Patent No. 5217870  
; APPLICANT: HESSTON, CATHERINE A.; LOBB, ROY R.; GOELZ, SUSAN E.  
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES AGAINST CDX  
; NUMBER OF SEQUENCES: 4  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/345,151  
; FILING DATE: 28-APR-1989  
; SEQ ID NO: 2  
; LENGTH: 610  
5217870-2

Query Match 60.5%; Score 130; DB 6; Length 610;  
Best Local Similarity 63.6%; Pred. No. 3e-07;  
Matches 21; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWCSGHEGCEVTINNTYCKDCDPGFSGLKCE 33  
Db 143 CTNTSCGHEGCEVTINNTYCKDCDPGFSGLKCE 175

RESULT 16  
US-08-340-539A-28  
; Sequence 28, Application US/08340539A  
; Patent No. 5808025  
; GENERAL INFORMATION:  
; APPLICANT: Tedder, Thomas F.  
; APPLICANT: Kansas, Geoffrey S.  
; TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
; NUMBER OF SEQUENCES: 28  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FISH & NEAVE  
; STREET: 1251 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10020  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/340,539A  
; FILING DATE: 16-NOV-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/008,459  
; FILING DATE: 25-JAN-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gunnison, Jane  
; REGISTRATION NUMBER: 38,479  
; REFERENCE/DOCKET NUMBER: CG-104 CON  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-596-9000  
; TELEFAX: 212-596-9090  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 36 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-340-539A-28

Query Match 59.1%; Score 127; DB 1; Length 36;  
Best Local Similarity 57.6%; Pred. No. 4e-08;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Oy 1 CQWMSGSGHCEVEIINHTCNCQDVGYYGPOCQ 33  
|| ||| |||:| |:||:| |:||:| |:  
Db 3 CODMCSKQGECEITIGNTGTCSCYGFYGPCE 35

RESULT 17  
US-08-340-539A-24  
Sequence 24, Application US/08340539A  
Patent No. 5808025  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340.539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-24

Query Match 58.6%; Score 126; DB 1; Length 36;  
Best Local Similarity 57.6%; Pred. No. 5.2e-08;  
Matches 19; Conservative 5; Mismatches 9; Indels 0; Gaps 0;  
Oy 1 CQWMSGSGHCEVEIINHTCNCQDVGYYGPOCQ 33  
|| ||| |||:| |:||:| |:||:| |:  
Db 3 CTNMSCSGHCEITINSTCKCHPFLGPNC 35

RESULT 18  
US-08-340-539A-26  
Sequence 26, Application US/08340539A  
Patent No. 5808025  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340.539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 26:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-26

Query Match 58.1%; Score 125; DB 1; Length 36;  
Best Local Similarity 57.6%; Pred. No. 6.7e-08;  
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0;

Oy 1 CQWMSGSGHCEVEIINHTCNCQDVGYYGPOCQ 33  
|| ||| |||:| |:||:| |:||:| |:  
Db 3 CODMCSKQGECEITIGNTGTCSCYGFYGPCE 35

RESULT 19  
US-08-252-493C-9  
Sequence 9, Application US/08252493C  
Patent No. 5891645  
GENERAL INFORMATION:  
APPLICANT: Rollins, Scott  
APPLICANT: Rother, Russell P.  
APPLICANT: Evans, Mark J.  
APPLICANT: Malis, Louis A.  
TITLE OF INVENTION: PORCINE E-SELECTIN  
NUMBER OF SEQUENCES: 9  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Seth A. Fidel  
STREET: 25 Science Park, Box 15  
CITY: New Haven  
STATE: Connecticut  
COUNTRY: USA  
ZIP: 06511  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 750 kb storage  
COMPUTER: PC compatible  
OPERATING SYSTEM: DOS 6.2  
SOFTWARE: WordPerfect 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/252.493C  
FILING DATE: June 1, 1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:



TITLE OF INVENTION: Peptides  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street, Suite 3100  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: US  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/05059  
FILING DATE: 19910717  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/320408  
FILING DATE: 08-MAR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/554199  
FILING DATE: 17-JUL-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRFL10CIP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6558  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 830 amino acids  
TYPE: AMINO ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: YES  
ANTI-SENSE: YES  
FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: Blood  
CELL TYPE: Endothelial  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 4..25  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 60..158  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 131..150  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 163..174  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 168..183  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 185..194  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 200..244  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 213..226  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 230..257  
NAME/KEY: Disulfide-bond  
LOCATION: 262..306  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 275..288  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 292..319  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 324..368  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 337..350  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 354..381  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 386..430  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 399..412  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 416..443  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 448..492  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 461..474  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 478..505  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 510..554  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 523..536  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 540..567  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 572..616  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 585..616  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 602..629  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 642..686  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 655..668  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 672..699  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 704..748  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 717..730  
FEATURE:  
NAME/KEY: Disulfide-bond  
LOCATION: 734..761  
FEATURE:  
NAME/KEY: Binding-site

```

Query Match          58.1%; Score 125; DB 6; Length 830.
Best Local Similarity 57.6%; Pred. No.1.5e-06;
Matches 19; Conservative 6; Mismatches 8; Indels 0; Gaps 0

Qy      1 CQPWSCSGHGECVEIINNHTCNCDDVGYGPOCQ 33
        ||| ||||| | : ||| | : ||| | : ||| | :
Db      163 CQDMSCSKGEGCELETIGNYTCSCYPFGPECE 195

RESULT 24
US-08-340-539A-25
; Sequence 25, Application US/08340539A
; Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSER: FISH & NEAVE
STREET: 1251 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10020
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340.539A
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008.459
FILING DATE: 25-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Gunnison, Jane
REGISTRATION NUMBER: 38,479
REFERENCE/DOCKET NUMBER: CG-104 CON
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-596-9000
TELEFAX: 212-596-9090
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 36 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539A-25

Query Match          57.2%; Score 123; DB 1; Length 36;
Best Local Similarity 57.6%; Pred. No.1.1e-07;
Matches 19; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

Qy      1 CQPWSCSGHGECVEIINNHTCNCDDVGYGPOCQ 33
        ||| ||||| | : ||| | : ||| | : ||| | :
Db      3 CNPTPCGSHGECVEIINNNTCCQHPFGKACE 35

RESULT 25
US-08-340-539A-27
; Sequence 27, Application US/08340539A
; Patent No. 5808025
GENERAL INFORMATION:
APPLICANT: Tedder, Thomas F.
APPLICANT: Kansas, Geoffrey S.
```

TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/340,539A  
FILING DATE: 16-NOV-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Gunnison, Jane  
REGISTRATION NUMBER: 38,479  
REFERENCE/DOCKET NUMBER: CG-104 CON  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000  
TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 27:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 36 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-340-539A-27

Query Match 55.8%; Score 120; DB 1; Length 36;  
Best Local Similarity 51.5%; Pred. No. 2.4e-07;  
Matches 17; Conservative 8; Mismatches 8; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCVDVGYGPQCQ 33  
Db 3 CODMSCSNGKCIETIGSYTSCYGFYGPCE 35

RESULT 26  
US-09-068-740A-3  
Sequence 3, Application US/09068740A  
Patent No. 6337387  
GENERAL INFORMATION:  
APPLICANT: SAKANO, SEIJI  
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE  
FILE REFERENCE: KP-8447  
CURRENT APPLICATION NUMBER: US/09/068,740A  
CURRENT FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: JP 7-299611  
PRIOR FILING DATE: 1995-11-17  
PRIOR APPLICATION NUMBER: JP 7-311811  
PRIOR FILING DATE: 1995-11-30  
PRIOR APPLICATION NUMBER: PCT/JP96/03356  
PRIOR FILING DATE: 1996-11-15  
SOFTWARE: Patentin Ver. 2.1  
NUMBER OF SEQ ID NOS: 48  
SEQ ID NO 3  
LENGTH: 520  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-09-068-740A-3

Query Match 43.3%; Score 93; DB 4; Length 520;  
Best Local Similarity 38.2%; Pred. No. 0.0036;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCVDVGYGPQCQ 34  
Db 311 CDPSPCKNGSCTDLENSYSCYCPGFYGCICEL 344

RESULT 27  
US-08-872-855-10  
Sequence 10, Application US/08872855  
Patent No. 6121045  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean  
APPLICANT: Gearity, David  
TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND  
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HONG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/872,855  
FILING DATE: 11-JUN-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MAA-003.02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-1000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 642 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-872-855-10

Query Match 43.3%; Score 93; DB 3; Length 642;  
Best Local Similarity 41.2%; Pred. No. 0.0044;  
Matches 14; Conservative 7; Mismatches 13; Indels 0; Gaps 0;

QY 1 COPWSCSGHCEVEIINNHHTCNCVDVGYGPQCQ 34  
Db 428 CASSPCANGCITCVDAVNSYTCSTLGYGKDCITL 461

RESULT 28  
US-09-068-740A-4  
Sequence 4, Application US/09068740A  
Patent No. 6337387  
GENERAL INFORMATION:  
APPLICANT: SAKANO, SEIJI  
TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE  
FILE REFERENCE: KP-8447  
CURRENT APPLICATION NUMBER: US/09/068,740A  
CURRENT FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: JP 7-299611

;; PRIOR FILING DATE: 1995-11-17  
;; PRIOR APPLICATION NUMBER: JP 7-311811  
;; PRIOR FILING DATE: 1995-11-30  
;; PRIOR APPLICATION NUMBER: PCT/JP96/03356  
;; PRIOR FILING DATE: 1996-11-15  
;; NUMBER OF SEQ ID NOS: 48  
;; SOFTWARE: Patentln Ver. 2.1  
;; SEQ ID NO: 4  
;; LENGTH: 702  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-09-068-740A-4

Query Match 43.3%; Score 93; DB 4; Length 702;  
Best Local Similarity 38.2%; Pred. No. 0.0048;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPMSCGHECEVEIINNHTCNCVGYGPOCQL 34  
DB 311 CDPSPCKNGSGCTDLENSYSCTCPFGYKICEL 344

RESULT 29  
US-09-068-740A-9  
; Sequence 9, Application US/09068740A  
; Patent No. 6337387  
; GENERAL INFORMATION:  
; APPLICANT: SAKANO, SEIJI  
; APPLICANT: ITOH, AKIRA  
; TITLE OF INVENTION: DIFFERENTIATION-SUPPRESSIVE POLYPEPTIDE  
; FILE REFERENCE: KP-8447  
; CURRENT APPLICATION NUMBER: US/09/068,740A  
; CURRENT FILING DATE: 1998-06-18  
; PRIOR APPLICATION NUMBER: JP 7-299611  
; PRIOR FILING DATE: 1995-11-17  
; PRIOR APPLICATION NUMBER: JP 7-311811  
; PRIOR FILING DATE: 1995-11-30  
; PRIOR APPLICATION NUMBER: PCT/JP96/03356  
; PRIOR FILING DATE: 1996-11-15  
; NUMBER OF SEQ ID NOS: 48  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO: 9  
; LENGTH: 723  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-068-740A-9

Query Match 43.3%; Score 93; DB 4; Length 723;  
Best Local Similarity 38.2%; Pred. No. 0.0049;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPMSCGHECEVEIINNHTCNCVGYGPOCQL 34  
DB 332 CDPSPCKNGSGCTDLENSYSCTCPFGYKICEL 365

RESULT 30  
US-08-185-432-18  
; Sequence 18, Application US/08185432  
; Patent No. 5750652  
; GENERAL INFORMATION:  
; APPLICANT: Aitavanis-Tsakonas, Spyridon  
; APPLICANT: Bussseau, Isabelle  
; APPLICANT: Diederich, Robert J.  
; APPLICANT: Xu, Tian  
; APPLICANT: Matsuno, Kenji  
; TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND  
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS

;; STREET: 1155 Avenue of the Americas  
;; CITY: New York  
;; STATE: New York  
;; COUNTRY: U.S.A.  
;; ZIP: 10036-2711  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: Patentln Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/185,432  
;; FILING DATE: 21-JAN-1994  
;; CLASSIFICATION: 530  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Mirock, S. Leslie  
;; REGISTRATION NUMBER: 18,872  
;; REFERENCE/DOCKET NUMBER: 7326-006  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: (212) 790-9090  
;; TELEFAX: (212) 869-8864/9741  
;; TELEX: 66141 PENNIE  
;; INFORMATION FOR SEQ ID NO: 18:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 2523 amino acids  
;; TYPE: amino acid  
;; TOPOLOGY: unknown  
US-08-185-432-18

Query Match 43.3%; Score 93; DB 1; Length 2523;  
Best Local Similarity 45.5%; Pred. No. 0.017;  
Matches 15; Conservative 4; Mismatches 14; Indels 0; Gaps 0;

OY 1 CQPMSCGHECEVEIINNHTCNCVGYGPOCQ 33  
DB 910 CQPNCHNGSGSDGIMFPCNCPAGFRGPKCE 942

RESULT 31  
US-08-872-855-7  
; Sequence 7, Application US/08872855  
; Patent No. 6121045  
; GENERAL INFORMATION:  
; APPLICANT: McCarthy, Sean  
; APPLICANT: Gearing, David  
; TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND  
; TITLE OF INVENTION: THERAPEUTIC USES THEREFOR  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
; STREET: One Post Office Square  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109-2170  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/872,855  
; FILING DATE: 11-JUN-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Arnold, Beth E.  
; REGISTRATION NUMBER: 35,430  
; REFERENCE/DOCKET NUMBER: MAA-003.02  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-832-1000  
; TELEFAX: 617-832-7000



; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 721 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-872-855-7

Query Match 41.9%; Score 90; DB 3; Length 721;  
Best Local Similarity 38.2%; Pred. No. 0.011;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHECEVETINNHNTGNCVDGYGPOCOL 34  
Db 334 CDANPCKNKGGSCSDLENSYTCSCPPGFYGNCEL 367

RESULT 32  
US-08-981-392-5  
; Sequence 5, Application US/08981392  
; Patent No. 6262025  
; GENERAL INFORMATION:  
; APPLICANT: Ish-Horowitz, David  
; APPLICANT: Henrique, Domingos Manuel Pinto  
; APPLICANT: Lewis, Julian Hart  
; APPLICANT: Artavanis-Tsakonas, Spyridon  
; APPLICANT: Gray, Grace  
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES  
; TITLE OF INVENTION: OF VERTEBRATE DELTA GENES AND METHODS BASED THEREON  
; NUMBER OF SEQUENCES: 94  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036/2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/981.392  
; FILING DATE: 22-DEC-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Antler, Adriane M.  
; REGISTRATION NUMBER: 32,605  
; REFERENCE/DOCKET NUMBER: 7326-038  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-8864  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 721 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-981-392-5

Query Match 41.9%; Score 90; DB 4; Length 721;  
Best Local Similarity 38.2%; Pred. No. 0.011;  
Matches 13; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHECEVETINNHNTGNCVDGYGPOCOL 34  
Db 334 CDANPCKNKGGSCSDLENSYTCSCPPGFYGNCEL 367

RESULT 33  
US-08-185-432-16  
; Sequence 16, Application US/08185432  
; Patent No. 5750652

; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, Spyridon  
; APPLICANT: Busseau, Isabelle  
; APPLICANT: Diederich, Robert J.  
; APPLICANT: Xu, Tian  
; APPLICANT: Matsuno, Kenji  
; TITLE OF INVENTION: DETEKT PROTEINS, NUCLEIC ACIDS, AND  
; TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PENNIE & EDMONDS  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/185.432  
; FILING DATE: 21-JAN-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 7326-006

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 16:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2471 amino acids  
; TYPE: amino acid  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-08-185-432-16

Query Match 41.4%; Score 89; DB 1; Length 2471;  
Best Local Similarity 41.2%; Pred. No. 0.046;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

Oy 1 COPWSCSGHECEVETINNHNTGNCVDGYGPOCOL 34  
Db 498 COSNPCVNNKGVNRFQCLCPGFTGPVQOI 531

RESULT 34  
US-08-083-590A-19  
; Sequence 19, Application US/08083590A  
; Patent No. 5786158  
; GENERAL INFORMATION:  
; APPLICANT: Artavanis-Tsakonas, S. et al.  
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods  
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.

ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/083,590A  
FILING DATE: 25-JUN-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7326-015  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 790-9090  
TELEFAX: 212 8698864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2471 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-083-590A-19

Query Match 41.4%; Score 89; DB 1; Length 2471;  
Best Local Similarity 41.2%; Pred. No. 0.046;  
Matches 14; Conservative 6; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVCEIINNHTCNCVGYGPOCOL 34  
DB 498 CQSNPCVNNGCYDKVNRFOCLCPGFTGVCQI 531

RESULT 35  
US-08-532-384-19  
Sequence 19, Application US/08512384  
Patent No. 6083904  
GENERAL INFORMATION:  
APPLICANT: Artavanis-Tsakonas, S. et al.  
TITLE OF INVENTION: Therapeutic And Diagnostic Methods  
TITLE OF INVENTION: And Compositions Based On No. 6083904ch Proteins And  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pennie & Edmonds  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/532,384  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/083,590  
FILING DATE: 25-JUN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7326-015  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212 790-9090  
TELEFAX: 212 8698864/9741

TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2471 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-532-384-19

Query Match 41.4%; Score 89; DB 3; Length 2471;  
Best Local Similarity 35.3%; Pred. No. 0.018;  
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVCEIINNHTCNCVGYGPOCOL 34  
DB 498 CQSNPCVNNGCYDKVNRFOCLCPGFTGVCQI 531

RESULT 36  
US-08-872-855-5  
Sequence 5, Application US/08872855  
Patent No. 6121045  
GENERAL INFORMATION:  
APPLICANT: McCarthy, Sean  
TITLE OF INVENTION: NOVEL HUMAN DELTA3 COMPOSITIONS AND  
TITLE OF INVENTION: THERAPEUTIC USES THEREFOR  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FOLEY, HOAG & ELIOT LLP  
STREET: One Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109-2170  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/872,855  
FILING DATE: 11-JUN-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Arnold, Beth E.  
REGISTRATION NUMBER: 35,430  
REFERENCE/DOCKET NUMBER: MAA-003.02  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-832-7000  
TELEFAX: 617-832-7000  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 713 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-872-855-5

Query Match 40.9%; Score 88; DB 3; Length 713;  
Best Local Similarity 35.3%; Pred. No. 0.018;  
Matches 12; Conservative 8; Mismatches 14; Indels 0; Gaps 0;

QY 1 COPWSCSGHGEVCEIINNHTCNCVGYGPOCOL 34  
DB 330 CAPSPCKNGSCTLDEDSYSCTCPGFTGKVCCL 363

RESULT 37



APPLICANT: Buseau, Isabelle  
APPLICANT: Diederich, Robert J.  
APPLICANT: Xu, Tian  
APPLICANT: Matsuno, Kenji  
TITLE OF INVENTION: DELTEX PROTEINS, NUCLEIC ACIDS, AND  
TITLE OF INVENTION: ANTIBODIES, AND RELATED METHODS AND COMPOSITIONS  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PENNIE & EDMONDS  
STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,432  
FILING DATE: 21-JAN-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 7326-006  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 790-9090  
TELEFAX: (212) 869-8864/9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 19:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2703 amino acids  
TYPE: amino acid  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-185-432-19

Query Match 40.0%; Score 86; DB 1; Length 2703;  
Best Local Similarity 42.4%; Pred. No. 0.11;  
Matches 14; Conservative 4; Mismatches 15; Indels 0; Gaps 0;  
QY 1 CQPMSCSGHGEVCEIINNHTCNCDVGYGPQCQ 33  
Db 1026 CTSSCLNGSCIDGNGCSCLAGISGANCO 1058

Search completed: September 7, 2002, 10:14:43  
Job time: 276 sec



```

XX 07-MAY-1996.
PD
XX
XX 23-FEB-1989; 89US-0315015.
PF
XX
XX 22-NOV-1989; 89US-0440625.
PR
XX 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GENE ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI; 1996-238773/24.
DR
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
XX Disclosure; Page 19; 41pp; English.
PS
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents; or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery. Improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is a Gly96-Ile97 deletion mutant.
CC
XX
XX Sequence 369 AA;
SQ
XX
XX Query Match 100.0%; Score 215; DB 17; Length 369;
XX Best Local Similarity 100.0%; Pred. No. 3e-14;
XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 COPWSCGHCVEEIIINHTCNCDDVGYGPQCQL 34
DB 157 CQPWSCGHCVEEIIINHTCNCDDVGYGPQCQL 190

```

```

FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..153
FT /label= Lectin domain.
FT Domain 158..191
FT /label= EGF domain.
FT Binding-site 195..256
FT /label= Complement binding repeat 1.
FT Binding-site 257..315
FT /label= Complement binding repeat 2.
FT Domain 331..353
FT /label= Transmembrane domain.
FT Domain 354..370
FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185670.
XX
XX (GENE ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI; 1996-238773/24.
DR
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
XX
XX Disclosure; Page 19; 41pp; English.
PS
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents; or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery. Improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is an Asn136 deletion mutant.
CC
XX
XX Sequence 370 AA;
SQ
XX
XX Query Match 100.0%; Score 215; DB 17; Length 370;
XX Best Local Similarity 100.0%; Pred. No. 3e-14;
XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 COPWSCGHCVEEIIINHTCNCDDVGYGPQCQL 34
DB 158 CQPWSCGHCVEEIIINHTCNCDDVGYGPQCQL 191

```

```
RESULT 3
AAR98128
ID AAR98128 standard; Protein: 370 AA.
XX
AC AAR98128;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain 38..154
FT Domain /label= Lectin domain.
FT Domain 159..192
FT Binding-site /label= EGF domain.
FT Binding-site 196..256
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site 257..315
FT Domain /label= Complement binding repeat 2.
FT Domain 331..353
FT Domain /label= Transmembrane domain.
FT Domain 354..370
FT Domain /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
```

```
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant is a Ser220 deletion mutant.
XX
SQ Sequence 370 AA;
XX
Query Match 100.0%; Score 215; DB 17; Length 370;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 COPWSCSGHGECVEIINHTCNCDDYGYGPOCQL 34
DB 159 cqpwscsghgceveIinhhtcncddvgyygpqcql 192
XX
RESULT 4
AAR98130
ID AAR98130 standard; Protein: 370 AA.
XX
AC AAR98130;
XX
DT 01-NOV-1996 (first entry)
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
KW metastasis.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain 38..154
FT Domain /label= Lectin domain.
FT Domain 159..192
FT Binding-site /label= EGF domain.
FT Binding-site 196..257
FT Binding-site /label= Complement binding repeat 1.
FT Binding-site 258..315
FT Binding-site /label= Complement binding repeat 2.
FT Domain /label= Complement binding repeat 2.
FT Domain 331..353
FT Domain /label= Transmembrane domain.
FT Domain 354..370
FT Domain /label= Cytoplasmic domain.
XX
PN US5514582-A.
XX
PD 07-MAY-1996.
XX
PF 23-FEB-1989; 89US-0315015.
XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
PI Capon DJ, Lasky LA;
XX
DR WPI; 1996-238773/24.
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
XX
PS Disclosure; Page 19; 41pp; English.
```

XX A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection, inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant is an Ile296 deletion mutant.  
 XX

Sequence 370 AA:

Query Match 100.0%; Score 215; DB 17; Length 370;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGEVEIINHTCNCNDVGYGPQCL 34  
 ||||||||||||||||||||||||||||  
 Db 159 cqpwsccshgceveiinhtcncndvgygpqcl 192

## RESULT 5

AAR98109  
 ID AAR98109 standard; Protein; 371 AA.

AC AAR98109;

DT 31-OCT-1996 (first entry)

XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 DE  
 XX  
 KW Immunoglobulin: transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 KW  
 XX  
 OS Homo sapiens.

XX  
 FH Key Location/Qualifiers  
 FT Region 1..37  
 FT /label= Signal region.  
 FT Domain 38..154  
 FT /label= Lectin domain.  
 FT Domain 159..192  
 FT /label= EGF domain.  
 FT Binding-site 196..257  
 FT /label= Complement binding repeat 1.  
 FT Binding-site 258..316  
 FT /label= Complement binding repeat 2.  
 FT Domain 332..354  
 FT /label= Transmembrane domain.  
 FT Domain 355..371  
 FT /label= Cytoplasmic domain.

XX US5514582-A.

XX 07-MAY-1996.

XX 23-FEB-1989; 89US-0315015.

XX

PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.

XX (GETH ) GENENTECH INC.

XX Capon DJ, Lasky LA;

XX WPI, 1996-238773/24.

XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation

XX Disclosure; Page 19; 41pp; English.

XX A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection, inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains a Arg58-Asp59; Lys-Glu  
 CC substitution.  
 XX

SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CQPMSCSGHGEVEIINHTCNCNDVGYGPQCL 34  
 ||||||||||||||||||||||||||||  
 Db 159 cqpwsccshgceveiinhtcncndvgygpqcl 192

## RESULT 6

AAR98110  
 ID AAR98110 standard; Protein; 371 AA.

AC AAR98110;

DT 01-NOV-1996 (first entry)

XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 DE  
 XX  
 KW Immunoglobulin: transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 KW  
 XX  
 OS Homo sapiens.

XX

XX Key Location/Qualifiers

XX Region 1..37  
 FT /label= Signal region.  
 FT Domain 38..154

FT



Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps
1 COPWSCSGHGECEVETIINNHTCNCNDVGYGPOCOL 34	100.0%; Score 215; DB 17; Length 371;	34	0	0	0	0
159 cqpwscsgshgecevetiinnhtcncndvgyypqcl 192	100.0%; Score 215; DB 17; Length 371;	34	0	0	0	0

AC	AA898111;	
XX		
DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.	
XX		
KM	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KM	immunomodulator; cell adhesion; graft rejection; inflammation;	
KW	metastasis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
XX	23-FEB-1989;	89US-0315015.
PF		
XX	22-NOV-1989;	89US-0440625.
XX	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808112.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GENTH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI, 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure: Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AA898106 are given in	
CC	AA898109-898135. This variant contains an Lys78Gln substitution.	

```
XX SQ Sequence 371 AA;
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CQPMSCSGHGECEVEIINNHTCNCDCVGYGPOCOL 34
Db 159 cqpmscsghgeceveinnhtcncdcvgygppcql 192
|||||
RESULT 8
AAR98112
ID AAR98112 standard; Protein; 371 AA.
XX AAR98112;
AC AAR98112;
XX 01-NOV-1996 (first entry)
DT
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Region 1..37
FT /label= Signal region.
FT Domain 38..154
FT /label= Lectin domain.
FT Domain 159..192
FT /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.
FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT /label= Transmembrane domain.
FT Domain 355..371
FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 16-DEC-1991; 91US-0808122.
XX
XX 08-DEC-1992; 92US-0986931.
XX
XX 21-JAN-1994; 94US-0185670.
XX
XX (GENET ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to its constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
```

```
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targetting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains an Asp116Glu substitution.
XX
XX SQ Sequence 371 AA;
Query Match 100.0%; Score 215; DB 17; Length 371;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 CQPMSCSGHGECEVEIINNHTCNCDCVGYGPOCOL 34
Db 159 cqpmscsghgeceveinnhtcncdcvgygppcql 192
|||||
RESULT 9
AAR98113
ID AAR98113 standard; Protein; 371 AA.
XX AAR98113;
AC AAR98113;
XX 01-NOV-1996 (first entry)
DT
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX Homo sapiens.
XX
XX FH Key Location/Qualifiers
XX FT Region 1..37
XX FT /label= Signal region.
XX FT Domain 38..154
XX FT /label= Lectin domain.
XX FT Domain 159..192
XX FT /label= EGF domain.
XX FT Binding-site 196..257
XX FT /label= Complement binding repeat 1.
XX FT Binding-site 258..316
XX FT /label= Complement binding repeat 2.
XX FT Domain 332..354
XX FT /label= Transmembrane domain.
XX FT Domain 355..371
XX FT /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 16-DEC-1991; 91US-0808122.
XX
XX 08-DEC-1992; 92US-0986931.
```

```

PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
PA
XX Capon DJ, Lasky LA;
PI
XX WPI: 1996-238773/24.
DR
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS
XX Disclosure; Page 19; 41pp; English.
XX
CC A hybrid immunoglobulin chain comprising the ligand binding site of
CC a single transmembrane receptor without an active transmembrane
CC region; fused at its C-terminus with the N-terminus of an
CC immunoglobulin constant region. The receptor is not a member of the
CC immunoglobulin super family, nor a multiple subunit polypeptide
CC encoded by discrete genes. The hybrid immunoglobulin chain combines
CC the adhesion/targeting of a ligand binding partner (LBP) with the
CC effector functions of immunoglobulin and can bind to and/or activate
CC more than one ligand. It can be used diagnostically for the in
CC vitro assay of LBP and their targets; or therapeutically to deliver
CC LBP such as toxins, enzymes, growth factors to particular cells.
CC Typical applications are as antiviral, neuromodulating and
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC The immunoglobulin component increases plasma half life and
CC facilitates purification while deletion of the transmembrane region
CC facilitates recovery, improves aqueous solubility and removes
CC potentially immunogenic epitopes. Variants of the human lymphocyte
CC cell surface glycoprotein described in AAR98106 are given in
CC AAR98109-R98135. This variant contains a Leu150Val substitution.
XX
XX Sequence 371 AA;
SQ
QY 1 COPWSCSGHGECEVEIINNHTCNCDDVGYGPOCOL 34
DB 159 cqpwsesghgecevelinnhtcncdvggygpqcl 192

RESULT 10
AAR98117
ID AAR98117 standard; Protein: 371 AA.
XX
XX AAR98117;
AC
XX 01-NOV-1996 (first entry)
DT
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX
XX Homo sapiens.
OS
XX
FH Key Location/Qualifiers
FT Region 1..37
FT Domain /label= Signal region.
FT Domain 38..154
FT Domain /label= Lectin domain.
FT Domain 159..192
FT Domain /label= EGF domain.
FT Binding-site 196..257
FT /label= Complement binding repeat 1.

```

```

FT Binding-site 258..316
FT /label= Complement binding repeat 2.
FT Domain 332..354
FT Domain /label= Transmembrane domain.
FT Domain 355..371
FT Domain /label= Cytoplasmic domain.
PN
XX US5514582-A.
XX
XX 07-MAY-1996.
PD
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
PI
XX WPI: 1996-238773/24.
DR
XX
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT binding site of a receptor fused to Ig constant region - useful for
PT diagnosis and treatment e.g. of inflammation
PS
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
XX facilitates recovery, improves aqueous solubility and removes
XX potentially immunogenic epitopes. Variants of the human lymphocyte
XX cell surface glycoprotein described in AAR98106 are given in
XX AAR98109-R98135. This variant contains a Thr151Ser substitution.
XX
XX Sequence 371 AA;
SQ
QY 1 COPWSCSGHGECEVEIINNHTCNCDDVGYGPOCOL 34
DB 159 cqpwsesghgecevelinnhtcncdvggygpqcl 192

RESULT 11
AAR98118
ID AAR98118 standard; Protein: 371 AA.
XX
XX AAR98118;
AC
XX 01-NOV-1996 (first entry)
DT
XX
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

```

```

XX      Immunoglobulin; transmembrane receptor; adhesion; targetting;
KM      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM      immunomodulator; cell adhesion; graft rejection; inflammation;
KM      metastasis;
XX      Homo sapiens.
OS
XX      Key
FH      Location/Qualifiers
FT      Region
FT      1..37
FT      /label= Signal region.
FT      Domain
FT      38..154
FT      /label= Lectin domain.
FT      Domain
FT      159..192
FT      /label= EGF domain.
FT      Binding-site
FT      196..257
FT      /label= Complement binding repeat 1.
FT      Binding-site
FT      258..316
FT      /label= Complement binding repeat 2.
FT      Domain
FT      332..354
FT      /label= Transmembrane domain.
FT      Domain
FT      355..371
FT      /label= Cytoplasmic domain.
XX
XX      US5514582-A.
PN
XX
XX      07-MAY-1996.
PD
XX
XX      23-FEB-1989; 89US-0315015.
PF
XX
XX      22-NOV-1989; 89US-0440625.
PR      23-FEB-1989; 89US-0315015.
PR      16-DEC-1991; 91US-0808122.
PR      08-DEC-1992; 92US-0986931.
PR      21-JAN-1994; 94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
PA
XX
XX      Capon DJ, Lasky LA;
PI
XX
XX      WPI; 1996-238773/24.
DR
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure; Page 19; 41pp; English.
PS
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targetting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate
CC      more than one ligand. It can be used diagnostically for the in
CC      vitro assay of LBP and their targets; or therapeutically to deliver
CC      LBP such as toxins, enzymes, growth factors to particular cells.
CC      Typical applications are as antiviral, neuromodulating and
CC      immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC      The immunoglobulin component increases plasma half life and
CC      facilitates purification while deletion of the transmembrane region
CC      facilitates recovery, improves aqueous solubility and removes
CC      potentially immunogenic epitopes. Variants of the human lymphocyte
CC      cell surface glycoprotein described in AAR98106 are given in
CC      AAR98109-R98135. This variant contains a Phe214Leu substitution.
XX
XX      Sequence 371 AA;
SQ
Query Match 100.0%; Score 215; DB 17; Length 371;

```

```

Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy      1 CQPMSCSGHGECEYITINHTCNCDVGYGPQCOL 34
Db      159 cqpwsctghgeceveitinhctncdvgygypqcl 192

RESULT 12
AAR98119
ID      AAR98119 standard; Protein, 371 AA.
XX
XX      AAR98119;
AC
XX
XX      01-NOV-1996 (first entry)
DT
XX
XX      Human lymphocyte cell surface glycoprotein (HuLHR) variant.
DE
XX      Immunoglobulin; transmembrane receptor; adhesion; targetting;
KM      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KM      immunomodulator; cell adhesion; graft rejection; inflammation;
KM      metastasis.
XX
XX      Homo sapiens.
OS
XX
XX      Key
FH      Location/Qualifiers
FT      Region
FT      1..37
FT      /label= Signal region.
FT      Domain
FT      38..154
FT      /label= Lectin domain.
FT      Domain
FT      159..192
FT      /label= EGF domain.
FT      Binding-site
FT      196..257
FT      /label= Complement binding repeat 1.
FT      Binding-site
FT      258..316
FT      /label= Complement binding repeat 2.
FT      Domain
FT      332..354
FT      /label= Transmembrane domain.
FT      Domain
FT      355..371
FT      /label= Cytoplasmic domain.
XX
XX      US5514582-A.
PN
XX
XX      07-MAY-1996.
PD
XX
XX      23-FEB-1989; 89US-0315015.
PF
XX
XX      22-NOV-1989; 89US-0440625.
PR      23-FEB-1989; 89US-0315015.
PR      16-DEC-1991; 91US-0808122.
PR      08-DEC-1992; 92US-0986931.
PR      21-JAN-1994; 94US-0185670.
XX
XX      (GETH ) GENENTECH INC.
PA
XX
XX      Capon DJ, Lasky LA;
PI
XX
XX      WPI; 1996-238773/24.
DR
XX
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX
XX      Disclosure; Page 19; 41pp; English.
PS
XX
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targetting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate

```

CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a Ser226Thr substitution.  
CC  
XX  
SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNDVGYGPGCOL 34  
Db 159 CQPMSCSGHGEVETIINHTCNDVGYGPGCOL 192  
|||||

RESULT 13  
AAR98120  
ID AAR98120 standard; Protein; 371 AA.  
XX  
AC AAR98120;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
DE  
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KM diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KM metastasis.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Region 1..37  
FT /label= Signal region.  
FT Domain 38..154  
FT /label= Lectin domain.  
FT Domain 159..192  
FT /label= EGF domain.  
FT Binding-site 196..257  
FT /label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /label= Complement binding repeat 2.  
FT Domain 332..354  
FT /label= Transmembrane domain.  
FT Domain 355..371  
FT /label= Cytoplasmic domain.  
XX  
XX US5514582-A.  
PN  
XX  
PD 07-MAY-1996.  
XX  
XX 23-FEB-1989; 89US-0315015.  
XX  
XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX Capon DJ, Lasky LA;  
PI

XX  
DR WPI: 1996-238773/24.  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX  
PS Disclosure: Page 19; 41pp; English.  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a Phe244Met substitution.  
CC  
XX  
SO Sequence 371 AA;

Query Match 100.0%; Score 215; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNDVGYGPGCOL 34  
Db 159 CQPMSCSGHGEVETIINHTCNDVGYGPGCOL 192  
|||||

RESULT 14  
AAR98121  
ID AAR98121 standard; Protein; 371 AA.  
XX  
AC AAR98121;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
DE  
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KM diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Region 1..37  
FT /label= Signal region.  
FT Domain 38..154  
FT /label= Lectin domain.  
FT Domain 159..192  
FT /label= EGF domain.  
FT Binding-site 196..257  
FT /label= Complement binding repeat 1.  
FT Binding-site 258..316  
FT /label= Complement binding repeat 2.  
FT Domain 332..354  
FT /label= Transmembrane domain.  
FT Domain 355..371  
FT

FT XX /label= Cytoplasmic domain.  
PN XX US5514582-A.  
XX XX  
PD 07-MAY-1996.  
XX XX  
PF 23-FEB-1989; 89US-0315015.  
XX XX  
PR 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX XX  
PA (GETH ) GENENTECH INC.  
XX XX  
PI Capon DJ, Lasky LA;  
XX XX  
DR WPI, 1996-238773/24.  
XX XX  
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX XX  
PS Disclosure; Page 19; 41pp; English.  
XX XX  
CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while detection of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a Thr282Ser substitution.  
XX XX  
SQ Sequence 371 AA:  
  
Query Match 100.0%; Score 215; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 CQPMSCSGHGECVEIINNHTCNCNDVGYGPPCQL 34  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
DB 159 CQPMSCSGHGECVEIINNHTCNCNDVGYGPPCQL 192  
  
RESULT 15  
AAR98122  
ID AAR98122 standard; Protein; 371 AA.  
XX XX  
AC AAR98122;  
XX XX  
DT 01-NOV-1996 (first entry)  
XX XX  
DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
XX XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.

XX XX Homo sapiens.  
OS XX  
XX XX Location/Qualifiers  
FH Key 1..37  
FT Region /label= Signal region.  
FT 38..154  
FT Domain /label= Lectin domain.  
FT 159..192  
FT Domain /label= EGF domain.  
FT 196..257  
FT Binding-site /label= Complement binding repeat 1.  
FT 258..316  
FT Binding-site /label= Complement binding repeat 2.  
FT 332..354  
FT Domain /label= Transmembrane domain.  
FT 355..371  
FT Domain /label= Cytoplasmic domain.  
XX XX  
XX XX US5514582-A.  
PN XX  
PD 07-MAY-1996.  
XX XX  
PF 23-FEB-1989; 89US-0315015.  
XX XX  
PR 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX XX  
PA (GETH ) GENENTECH INC.  
XX XX  
PI Capon DJ, Lasky LA;  
XX XX  
DR WPI, 1996-238773/24.  
XX XX  
PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
XX XX  
PS Disclosure; Page 19; 41pp; English.  
XX XX  
CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while detection of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains an Ile288Val substitution.  
XX XX  
SQ Sequence 371 AA:  
  
Query Match 100.0%; Score 215; DB 17; Length 371;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
OY 1 CQPMSCSGHGECVEIINNHTCNCNDVGYGPPCQL 34  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

Db 159 cqpwscsghecvellmhtncdvygypqcq1 192

RESULT 16

AAR98123

ID AAR98123 standard; Protein: 371 AA.

AC AAR98123;

XX

DT 01-NOV-1996 (first entry)

XX

DE Human lymphocyte cell surface glycoprotein (HLHR) variant.

XX

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

XX

OS Homo sapiens.

XX

PH Key

FT Region

FT Domain

FT Domain

FT Domain

FT Binding-site

FT Binding-site

FT Binding-site

FT Domain

FT Domain

FT Domain

XX

PN US5514582-A.

PD 07-MAY-1996.

XX

PF 23-FEB-1989; 89US-0315015.

XX

PR 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX

PA (GETH ) GENENTECH INC.

PI Capon DJ, Lasky LA;

XX

DR WPI; 1996-238773/24.

XX

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand

PT binding site of a receptor fused to Ig constant region - useful for

PT diagnosis and treatment e.g. of inflammation

XX

PS Disclosure; Page 19; 41pp; English.

XX

CC A hybrid immunoglobulin chain comprising the ligand binding site of

CC a single transmembrane receptor without an active transmembrane

CC region; fused at its C-terminus with the N-terminus of an

CC immunoglobulin constant region. The receptor is not a member of the

CC immunoglobulin super family, nor a multiple subunit polypeptide

CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the

CC effector functions of immunoglobulin and can bind to and/or activate

CC more than one ligand. It can be used diagnostically for the in

CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.

CC Typical applications are as antiviral, neuromodulating and

CC immunomodulating agents, or as modulators of cell adhesion (e.g. in

CC treating graft rejection; inflammation; metastasis of lymphoma etc.)

CC The immunoglobulin component increases plasma half life and

CC facilitates purification while deletion of the transmembrane region

CC facilitates recovery, improves aqueous solubility and removes

CC potentially immunogenic epitopes. Variants of the human lymphocyte

CC cell surface glycoprotein described in AAR98106 are given in

CC AAR98109-R98135. This variant contains a Lys298-Lys299; Arg-Arg

CC substitution.

XX

XX Sequence 371 AA;

SQ

Query Match 100.0%; Score 215; DB 17; Length 371;

Best Local Similarity 100.0%; Pred. No. 3e-14;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CQPWSCSGHECVELLMHTNCBDVGYGPOCOL 34

|||||

Db 159 cqpwscsghecvellmhtncdvygypqcq1 192

RESULT 17

AAR98124

ID AAR98124 standard; Protein: 371 AA.

XX

AC AAR98124;

XX

DT 01-NOV-1996 (first entry)

XX

DE Human lymphocyte cell surface glycoprotein (HLHR) variant.

XX

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;

KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;

KW immunomodulator; cell adhesion; graft rejection; inflammation;

KW metastasis.

XX

OS Homo sapiens.

XX

PH Key

FT Region

FT Domain

FT Domain

FT Domain

FT Binding-site

FT Binding-site

FT Binding-site

FT Domain

FT Domain

FT Domain

XX

PN US5514582-A.

PD 07-MAY-1996.

XX

PF 23-FEB-1989; 89US-0315015.

XX

PR 22-NOV-1989; 89US-0440625.

PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.

PR 08-DEC-1992; 92US-0986931.

PR 21-JAN-1994; 94US-0185670.

XX

PA (GETH ) GENENTECH INC.

PI Capon DJ, Lasky LA;

XX

DR WPI; 1996-238773/24.

XX

PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand





FT Protein 38..372  
 FT /label= probable mature protein  
 XX WO9108298-A.  
 XX 13-JUN-1991.  
 XX 21-NOV-1990; 90WO-US06849.  
 XX 22-NOV-1989; 89US-0444625.  
 XX (GETH ) GENENTECH INC.  
 XX Capon DJ, Lasky LA;  
 XX WPI; 1991-193202/26.  
 XX N-PSDB; AAQ12118.  
 DR New hybrid immunoglobulin(s) - for use as diagnostic reagents for  
 PT ligand binding molecules and to treat organ and graft rejection  
 PT and inflammation.  
 XX Disclosure: Fig 1; 67pp; English.  
 XX The gene product may be used as a ligand binding partner in combina-  
 CC tion with a stable plasma protein eg. IgG1-IgG4, IgA, IgE, IgD or IgM.  
 CC The fusion product is joined by N- or C-terminal groups, preferably  
 CC the N-terminal of the Fc region of the spp is linked to the C-terminal  
 CC of lbp. They may be used to provide antiviral, immunomodulatory  
 CC and neuroendocrine treatment as well as in treatment of receptor  
 CC mediated abnormalities.  
 XX Sequence 372 AA;  
 SQ  
 Query Match 100.0%; Score 215; DB 12; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 COPWSCSGHCEVETIINHTCNCDDVGYGPOCOL 34  
 Db 160 cqpwsctgsghevcvcllnhtcncdvygygpcq1 193  
 RESULT 20  
 AAR22802  
 ID AAR22802 standard; Protein; 372 AA.  
 XX  
 AC AAR22802;  
 XX  
 DT 01-SEP-1992 (first entry)  
 XX  
 DE Human lymphocyte homing receptor.  
 XX  
 KW HuLHR; LHR; binding; endothelium; immunogens; graft; organ;  
 KM rejection; inflammation; rheumatoid arthritis; lymphoma metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Region 20..32  
 FT /note= "potential signal sequence"  
 FT Modified-site 60..62 "N-glycosylation site"  
 FT Modified-site 104..106 "N-glycosylation site"  
 FT Modified-site 177..179 "N-glycosylation site"  
 FT Modified-site 216..218 "N-glycosylation site"  
 FT Modified-site 232..234 "N-glycosylation site"  
 FT Modified-site 271..273 "N-glycosylation site"

FT /note= "N-glycosylation site"  
 FT Modified-site 311..313  
 FT /note= "N-glycosylation site"  
 FT Region 335..357  
 FT /note= "stop transfer or membrane anchoring domain"  
 XX US5098833-A.  
 XX 24-MAR-1992.  
 XX 23-FEB-1989; 89US-0315015.  
 XX 23-FEB-1989; 89US-0315015.  
 XX (REGC ) UNIV OF CALIFORNIA.  
 XX Lasky LA, Rosen SD, Stachel SE, Singer MS, Yednock TA;  
 XX WPI; 1992-123385/15.  
 XX N-PSDB; AAQ23623.  
 DR New DNA encoding at least one domain of lymphocyte homing  
 PT receptor - useful for treating graft rejection, inflammation,  
 PT etc.  
 XX Disclosure: Fig 1; 32pp; English.  
 XX The protein sequence was deduced from the DNA sequence obtd. by  
 CC screening an oligo dt primed lambda gt10 cDNA library derived from  
 CC human peripheral blood lymphocyte mRNA obtd. from primary cells,  
 CC with a 2.2 kb EcoRI insert of the murine Mel 14 antigen cDNA clone.  
 CC The protein contains regions encoding a carbohydrate binding domain,  
 CC an EGF-like domain, a complement binding domain and a transmembrane  
 CC domain. The protein contains 26 cysteine residues. Cells transformed  
 CC by the hLHR DNA are used to produce LHR (which mediates binding of  
 CC lymphocytes to the endothelium of lymphoid tissue). LHR or its  
 CC variants are useful as reagents for assaying LHR or anti-LHR anti-  
 CC bodies, to purify the antibodies, as immunogens, and therapeutically  
 CC to compete with normal binding of lymphocytes (to prevent graft/organ  
 CC rejection; to treat inflammation (such as rheumatoid arthritis or  
 CC other autoimmune diseases); for control of lymphocyte metastasis,  
 CC and to treat conditions associated with accumulation of lymphocytes).  
 CC Derivs. and variants of LHR may be produced having modified  
 CC properties, e.g. increased activity, longer plasma half-life,  
 CC reduced side effects and better aq. solubility.  
 CC See also AAR22803.  
 XX  
 SQ Sequence 372 AA;  
 Query Match 100.0%; Score 215; DB 13; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 COPWSCSGHCEVETIINHTCNCDDVGYGPOCOL 34  
 Db 160 cqpwsctgsghevcvcllnhtcncdvygygpcq1 193  
 RESULT 21  
 AAR24026  
 ID AAR24026 standard; Protein; 372 AA.  
 XX  
 AC AAR24026;  
 XX  
 DT 22-NOV-1992 (first entry)  
 XX  
 DE Sequence of human lymphocyte cell surface glycoprotein  
 DE (HuLHR).  
 XX  
 KW Lymphocyte cell surface glycoprotein; ligand binding protein.  
 XX  
 OS Homo sapiens.

```

XX Key Location/Qualifiers
FH Peptide 1..19
FT Peptide 20..38
FT Peptide /label= signal
FT Modified-site 60..62
FT Modified-site /label= potential N-linked glycosylation site
FT Modified-site 104..106
FT Modified-site /label= see above
FT Modified-site 177..179
FT Modified-site /label= see above
FT Modified-site 216..218
FT Modified-site /label= see above
FT Modified-site 232..234
FT Modified-site /label= see above
FT Modified-site 271..273
FT Modified-site /label= see above
FT Modified-site 311..313
FT Modified-site /label= see above
FT Region 333..355
FT Region /label= stop transfer sequence
XX US5116964-A.
XX 26-MAY-1992.
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 22-NOV-1989; 89US-0440625.
XX (GETH ) GENENTECH INC.
XX Capon DJ, Lasky LA;
XX WPI: 1992-199589/24.
XX N-PSDB; AAQ24987.
XX Nucleic acid encoding polypeptide fusions - comprising ligand
XX binding partner protein and immunoglobulin chain, for use in
XX diagnosis and therapy
XX Disclosure; Fig 1-1 - 1-3; 43pp; English.
XX LHR mediates the binding of lymphocytes to the endothelium of
XX lymphoid tissue. Full length cDNA clones and DNA encoding the human
XX and the murine LHR (HulHR and MLHR, respectively) have been
XX identified and isolated (see AAQ24987 and AAQ24988). LHR is a
XX glycoprotein which contains the following protein domains: a signal
XX sequence, a carbohydrate binding domain, and epidermal growth
XX factor-like (egf) domain, at least one and preferably two complement
XX binding domain repeat, a transmembrane binding domain (TMD), and a
XX charged intracellular or cytoplasmic domain. LHR is used as the
XX ligand-binding partner in fusion polypeptides with an immunoglobulin,
XX for use in diagnosis and therapy.
XX
XX Sequence 372 AA:
XX
XX Query Match 100.0%; Score 215; DB 13; Length 372;
XX Best local similarity 100.0%; Pred. No. 3e-14;
XX Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 CQPMSCSGHGECEVEIINNHTCNCNCGVYGPQCOL 34
XX ||||||||||||||||||||||||||||||||||
XX Db 160 cgpwscsgbhgcveveinnhtcncnccvgygpqcol 193
XX
XX RESULT 22
XX AAR37960
XX ID AAR37960 standard; Protein; 372 AA.
XX AC AAR37960;

```

```

XX 08-OCT-1993 (first entry)
XX Human Lymphocyte Homing Receptor.
XX HulHR; Lymphocyte binding inhibition; lymphoma metastasis;
XX transplant rejection; inflammation.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH Peptide 1..38
FT Peptide /label= signal_sequence
FT Protein 39..372
FT Domain /note= "Trp39 is probable N-terminus of mature LHR"
FT Modified-site 39..155
FT Modified-site /label= Lectin_domain
FT Modified-site 60..62
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 104..106
FT Modified-site /note= "potential N-glycosylation site"
FT Domain 160..193
FT Domain /label= EGF_domain
FT Modified-site 177..179
FT Region /note= "potential N-glycosylation site"
FT Region 197..258
FT Modified-site /label= Complement_Binding_Repeat_1
FT Modified-site 216..218
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 232..234
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 246..248
FT Modified-site /note= "potential N-glycosylation site"
FT Region 259..317
FT Region /label= Complement_Binding_Repeat_2
FT Modified-site 271..273
FT Modified-site /note= "potential N-glycosylation site"
FT Modified-site 311..313
FT Modified-site /note= "potential N-glycosylation site"
FT Domain 333..355
FT Domain /label= Transmembrane_Domain
FT Domain /note= "stop transfer sequence"
FT Domain 356..372
FT Domain /label= Cytoplasmic_Domain
XX
XX US5216131-A.
XX 01-JUN-1993.
XX 23-FEB-1989; 89US-0315015.
XX 23-FEB-1989; 89US-0315015.
XX 23-FEB-1989; 89US-0315015.
XX 31-OCT-1991; 91US-0786149.
XX (GETH ) GENENTECH INC.
XX
XX Lasky LA, Rosen SD, Singer MS, Stachel SE, Yednock TA;
XX WPI: 1993-188588/23.
XX N-PSDB; AAQ43154.
XX Human and murine lymphocyte homing receptors to treat graft
XX rejection and inflammation - comprise carbohydrate binding,
XX epidermal growth factor and complement binding domains
XX
XX Claim 1; Fig 1 and Fig 3; 32pp; English.
XX
XX A human peripheral blood lymphocyte cDNA library in lambda gt10 was
XX screened with a 2.2kb EcoRI insert of the murine Mcl14 antigen clone
XX (i.e. a murine LHR sequence). The largest EcoRI insert (2.2kb) was
XX isolated and sequenced. The ORF codes for 372 amino acids with a mol.
XX wt. of approximately 42,200. Comparison of the HulHR amino acid
XX sequence with the murine LHR sequence (AAR37961) showed a high degree

```



```

FT FT /note= "putative stop transfer or membrane anchor
FT FT domain"
FT FT Region 358..372
FT FT /note= "putative intracellular region"
XX XX
PN PN US5428130-A.
XX XX
PD PD 27-JUN-1995.
XX XX
XX XX 23-FEB-1989; 89US-0315015.
XX XX
PR PR 22-NOV-1989; 89US-0440625.
PR PR 23-FEB-1989; 89US-0315015.
PR PR 16-DEC-1991; 91US-0808122.
PR PR 08-DEC-1992; 92US-0986931.
XX XX
PA PA (GETH ) GENENTECH INC.
XX XX
PI PI Capon DJ, Lasky LA;
XX XX
DR DR WPI; 1995-240086/31.
DR DR N-PSDB; AAO92802.
XX XX
PT PT New hybrid ligand binding partner molecules - fused to immunoglobulin
PT PT constant region sequences to increase stability and in vivo plasma
PT PT half-life
XX XX
PS PS Disclosure; Fig.1a-1c; 40pp; English.
XX XX
CC CC A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10
CC CC cDNA library derived from human peripheral blood lymphocyte mRNA
CC CC obtd. from primary cells. A cDNA clone encoding LHR was isolated.
XX XX
SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVETIINNHTCNCDDVGYGPQCQL 34
Db 160 cqpwsctgngceveilnnhtcncdvygypqcql 193

RESULT 25
AAR83050
ID AAR83050 standard; protein; 372 AA.
XX XX
AC AAR83050;
XX XX
DT 31-JAN-1996 (first entry)
XX XX
DE Human LHR.
XX XX
KW Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;
KW immunoglobulin; IgG; constant region; receptor-mediated disease;
KW vector; plasma-life.
XX XX
OS Homo sapiens.
XX XX
FH Key Location/Qualifiers
FH FT 20..32
FH FT /label= sig_peptide
FH FT 39..155
FH FT /label= Carbohydrate_binding_domain
FH FT 160..193
FH FT /label= Epidermal_growth_factor_domain
FH FT 197..317
FH FT /label= Complement_factor_binding_domain
FH FT 333..355
FH FT /label= Transmembrane_binding_domain
FH FT 356..372
FH FT Domain

```

```

FT FT /label= Cytoplasmic_domain
FT FT /note= "potential stop transfer sequence"
FT FT 60..62
FT FT /label= N-glycosylation_site
FT FT 104..106
FT FT /label= N-glycosylation_site
FT FT 177..179
FT FT /label= N-glycosylation_site
FT FT 216..218
FT FT /label= N-glycosylation_site
FT FT 232..234
FT FT /label= N-glycosylation_site
FT FT 248..248
FT FT /label= N-glycosylation_site
FT FT 271..273
FT FT /label= N-glycosylation_site
FT FT 311..313
FT FT /label= N-glycosylation_site
XX XX
PN PN US5455165-A.
XX XX
PD 03-OCT-1995.
XX XX
PE 23-FEB-1989; 89US-0315015.
XX XX
PR 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185569.
XX XX
PA (GETH ) GENENTECH INC.
XX XX
PI Capon DJ, Lasky LA;
XX XX
DR WPI; 1995-350776/45.
DR DR N-PSDB; AAT05869.
XX XX
PT Expression vector encoding fusion protein to increase plasma life -
PT PT complex receptor ligand binding site and Ig constant region, for
PT PT treatment of receptor mediated disease
XX XX
PS Disclosure; Fig 1; 42pp; English.
XX XX
CC A mouse LHR (lymphocyte cell surface glycoprotein) cDNA clone was used
CC CC to screen an oligo-dt primed lambda gt10 cDNA library derived from
CC CC human peripheral blood lymphocyte mRNA obtd. from primary cells. A
CC CC 2.2 Kb clone (sequence given in AAT05869) was isolated that encoded the
CC CC human LHR protein (AAR83050). LHR-IgG hybrids were constructed for use
CC CC in the targeting of therapeutic moieties to lymphoid tissue.
XX XX
SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 3e-14;
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVETIINNHTCNCDDVGYGPQCQL 34
Db 160 cqpwsctgngceveilnnhtcncdvygypqcql 193

RESULT 26
AAR98106
ID AAR98106 standard; Protein; 372 AA.
XX XX
AC AAR98106;
XX XX
DT 31-OCT-1996 (first entry)
XX XX
DE Human lymphocyte cell surface glycoprotein (HuLHR).
XX XX

```

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..38  
 FT /label= Signal region.  
 FT Domain 39..155  
 FT /label= Lectin domain.  
 FT Domain 160..193  
 FT /label= EGF domain.  
 FT Binding-site 197..258  
 FT /label= Complement binding repeat 1.  
 FT Binding-site 259..317  
 FT /label= Complement binding repeat 2.  
 FT Domain 333..355  
 FT /label= Transmembrane domain.  
 FT Domain 356..372  
 FT /label= Cytoplasmic domain.  
 XX  
 PN US514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 XX  
 PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI: 1996-238773/24.  
 DR N-PSDB; AAR98106.  
 XX  
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation  
 XX  
 PS Example 2; Figure 1; 41pp; English.  
 XX  
 CC A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes.  
 CC  
 XX Sequence 372 AA;  
 XX

QY 1 CQWSCSGHCEVEITNNHTCNDGVGYGPQCL 34  
 ||||||||||||||||||||||||||||||||  
 DB 160 cqpwschgceveitnnhtcndvgygppqcl 193  
 RESULT 27  
 AAR98133  
 ID AAR98133 standard; Protein; 372 AA.  
 XX  
 AC AAR98133;  
 XX  
 DT 01-NOV-1996 (first entry)  
 XX  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 XX  
 XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..37  
 FT /label= Signal region.  
 FT Domain 38..154  
 FT /label= Lectin domain.  
 FT Domain 159..192  
 FT /label= EGF domain.  
 FT Binding-site 196..258  
 FT /label= Complement binding repeat 1.  
 FT Binding-site 259..317  
 FT /label= Complement binding repeat 2.  
 FT Domain 333..355  
 FT /label= Transmembrane domain.  
 FT Domain 356..372  
 FT /label= Cytoplasmic domain.  
 XX  
 PN US514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 XX  
 PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI: 1996-238773/24.  
 XX  
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation  
 XX  
 PS Disclosure: Page 19; 41pp; English.  
 XX  
 CC A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver

CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains a 209-Aan insertion.  
 XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 17; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVEIINHTCNCDDVGYGPQCL 34  
 |||  
 DB 159 cqpwsctghgeveiiinhctncddvgygpqcl 192

RESULT 28  
 AAW37781  
 ID AAW37781 standard; Protein; 372 AA.

XX AAW37781;

DT 17-AUG-1998 (first entry)

XX Homo sapiens lymphocyte homing receptor (LHR).

KW lymphocyte homing receptor; LHR; HULHR; organ; graft; rejection;  
 KW treatment; inflammatory disorders; rheumatoid arthritis;  
 KW autoimmune diseases; lymphoma metastasis; control; lymphocyte;  
 KW accumulation.  
 XX  
 OS Homo sapiens.

XX

XX Location/Qualifiers

FT Key 20..33  
 FT Peptide /note= "potential signal sequence"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

FT /note= "potential N-linked glycosylation site"

PR 26-MAY-1995; 9505-0451848.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI; 1998-129805/12.  
 XX  
 DR N-PSDB; AAV19012.  
 XX  
 PT Prevention of lymphocyte attachment to endothelial cells - using  
 PT chimeric molecule comprising lymphocyte homing receptor and  
 PT immunoglobulin constant region  
 XX  
 PS Disclosure; Fig 1; 43pp; English.  
 XX  
 CC The sequence is that of a human lymphocyte homing receptor  
 CC (LHR) which may be used in the construction of a chimeric molecule  
 CC comprising an LHR fused at its C terminus to the N terminus of an  
 CC immunoglobulin constant region. This can be used for the prevention  
 CC of lymphocyte attachment to endothelial cells. Such a method may  
 CC be used for preventing organ or graft rejection, for treating  
 CC inflammatory disorders, e.g. rheumatoid arthritis or other  
 CC autoimmune diseases, for controlling lymphoma metastasis and  
 CC for treating conditions in which there is an accumulation of  
 CC lymphocytes.  
 XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 215; DB 19; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 3e-14;  
 Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 COPWSCSGHGEVEIINHTCNCDDVGYGPQCL 34  
 |||  
 DB 160 cqpwsctghgeveiiinhctncddvgygpqcl 193

RESULT 29

AAW73264  
 ID AAW73264 standard; Protein; 372 AA.

XX AAW73264;

DT 02-FEB-1999 (first entry)

XX Human lymphocyte homing receptor.

KW lymphocyte homing receptor; LHR; lymphocyte cell-surface glycoprotein;  
 KW lymphocyte binding; endothelium; graft rejection; inflammation; therapy;  
 KW arthritis; autoimmune disease; lymphoma metastasis;  
 KW lymphocyte accumulation; human.  
 XX  
 OS Homo sapiens.

XX US5840844-A.  
 PN 24-NOV-1998.

XX 10-AUG-1995; 9505-0513278.

XX 23-FEB-1989; 8905-0315015.

XX 31-OCT-1991; 9105-0786149.

XX 06-MAY-1993; 9305-0059029.

XX 10-AUG-1995; 9505-0513278.

XX (GETH ) GENENTECH INC.  
 PA (REGC ) UNIV CALIFORNIA.

XX Lasky LA, Rosen SD, Slinger MS, Stachel SE;  
 PI WPI; 1999-034122/03.

XX N-PSDB; AAV08321.

XX

XX

XX

XX

XX

XX

XX  
PT Lymphocyte homing receptor polypeptides - useful for inhibiting  
PT Lymphocyte binding to lymphoid endothelium  
XX  
PS Claim 1; Fig 1; 33pp; English.  
XX  
CC This sequence is the human lymphocyte homing receptor (LHR) of the  
CC invention. LHR is a lymphocyte cell-surface glycoprotein that mediates  
CC the binding of lymphocytes to the endothelium of lymphoid tissue. Soluble  
CC LHR polypeptides, lacking signal peptide (amino acids 1-38),  
CC transmembrane domain (amino acids 333-355) and cytoplasmic domain (amino  
CC acids 356-372), can be used therapeutically to compete with the normal  
CC binding of lymphocytes to lymphoid tissue and are especially useful for  
CC organ or graft rejection treatment protocols, for treating inflammations  
CC such as arthritis and other autoimmune diseases, for control of lymphoma  
CC metastasis and for treating conditions involving lymphocyte accumulation.  
CC LHR polypeptides can also be used in assays for LHR, anti-LHR antibodies  
CC or competitive inhibitors of LHR activity, and for purifying anti-LHR  
CC antibodies, and as immunogens for raising anti-LHR antibodies.  
XX  
SO Sequence 372 AA;

Query Match 100.0%; Score 215; DB 20; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNDVGYGPQCQL 34  
Db 160 CQPMSCSGHGEVETIINHTCNDVGYGPQCQL 193

RESULT 30  
AAR98131  
ID AAR98131 standard; Protein; 374 AA.  
XX  
AC AAR98131;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
KM  
XX Homo sapiens.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT Region 1..37  
FT /label= Signal region.  
FT Domain 38..157  
FT /label= Lectin domain.  
FT Domain 162..195  
FT /label= EGF domain.  
FT Binding-site 199..260  
FT /label= Complement binding repeat 1.  
FT Binding-site 261..319  
FT /label= Complement binding repeat 2.  
FT Domain 335..357  
FT /label= Transmembrane domain.  
FT Domain 358..374  
FT /label= Cytoplasmic domain.

US514582-A.  
XX  
PN 07-MAY-1996.  
XX  
PD 23-FEB-1989; 89US-0315015.  
XX  
PF 22-NOV-1989; 89US-0440625.  
XX  
PR 23-FEB-1989; 89US-0315015.

PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX  
PA (GETH ) GENENTECH INC.  
XX  
PI Capon DJ, Lasky LA;  
XX  
DR WPL 1996-238773/24.  
XX  
PR Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PR binding site of a receptor fused to Ig constant region - useful for  
PR diagnosis and treatment e.g. of inflammation  
XX  
PS Disclosure: Page 19; 41pp; English.

XX  
CC A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets; or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a 67-Glu-Ser-Ala insertion.  
XX  
XX  
SO Sequence 374 AA;

Query Match 100.0%; Score 215; DB 17; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CQPMSCSGHGEVETIINHTCNDVGYGPQCQL 34  
Db 162 CQPMSCSGHGEVETIINHTCNDVGYGPQCQL 195

RESULT 31  
AAR98132  
ID AAR98132 standard; Protein; 374 AA.  
XX  
AC AAR98132;  
XX  
DT 01-NOV-1996 (first entry)  
XX  
DE Human lymphocyte cell surface glycoprotein (HuLHR) variant.  
XX  
KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
KM  
XX Homo sapiens.  
XX  
OS  
XX  
FH Key Location/Qualifiers  
FT Region 1..37  
FT /label= Signal region.  
FT Domain 38..157  
FT /label= Lectin domain.  
FT Domain 162..195  
FT /label= EGF domain.

FT	Blinding-site	199..260
FT	/label= Complement binding repeat 1.	
FT	Binding-site	261..319.
FT	/label= Complement binding repeat 2.	
FT	Domain	335..357
FT	/label= Transmembrane domain.	
FT	Domain	358..374
XX	/label= Cytoplasmic domain.	
PX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
PA	(GETH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
XX	WPI; 1996-238773/24.	
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
XX	Disclosure: Page 19; 41pp; English.	
CC		
CC	A hybrid immunoglobulin chain comprising the ligand binding site, of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents; or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery. Improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-R98115. This variant contains a 83-Gly-Thr insertion.	
XX		
SQ	Sequence 374 AA:	
OY	Query Match 100.0%; Score 215; DB 17; Length 374;	
	Best Local Similarity 100.0%; Pred. No. 3.Le-1a;	
	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
DB	1 COPWSCGHCCEVEIINHTNCMDGYGYPQOL 34           162 cqpwschgceveeiinhtncmdvgygypqql 195	
RESULT 32		
ID AAR98134		
AC AAR98134 standard; Protein; 374 AA.		
XX		
XX AAR98134;		
DT 01-NOV-1996 (first entry)		

XX	Human lymphocyte cell surface glycoprotein (HuLHR) variant.
DE	
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KM	metastasis.
OS	
XX	Homo sapiens.
XX	
FH	
FT	Key
FT	Region
FT	/label= Signal region.
FT	38..154
FT	/label= Lectin domain.
FT	159..192
FT	/label= EGF domain.
FT	196..260
FT	/label= Complement binding repeat 1.
FT	261..319
FT	/label= Complement binding repeat 2.
FT	335..357
FT	/label= Transmembrane domain.
FT	358..374
FT	/label= Cytoplasmic domain.
PN	
PN	US514582-A.
PD	
XX	
XX	07-MAY-1996.
PF	
PR	23-FEB-1989; 89US-0315015.
PR	
PR	22-NOV-1989; 89US-0440625.
PR	23-FEB-1989; 89US-0315015.
PR	16-DEC-1991; 91US-0808122.
PR	08-DEC-1992; 92US-0986931.
PR	21-JAN-1994; 94US-0185670.
PA	(GERT ) GENENTECH INC.
P1	
PI	Capon DJ, Lasky LA;
DR	
XX	WPI; 1996-238773/24.
PT	
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT	binding site of a receptor fused to Ig constant region - useful for
PT	diagnosis and treatment e.g. of inflammation
PS	
PS	Disclosure; Page 19; 41pp; English.
CC	
CC	A hybrid immunoglobulin chain comprising the ligand binding site of
CC	a single transmembrane receptor without an active transmembrane
CC	region, fused at its C-terminus with the N-terminus of an
CC	immunoglobulin constant region. The receptor is not a member of the
CC	immunoglobulin super family, nor a multiple subunit polypeptide
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines
CC	the adhesion/targeting of a ligand binding partner (LBP) with the
CC	effector functions of immunoglobulin and can bind to and/or activate
CC	more than one ligand. It can be used diagnostically for the in
CC	viro assay of LBP and their targets, or therapeutically to deliver
CC	LBP such as toxins, enzymes, growth factors to particular cells.
CC	Typical applications are as antiviral, neuromodulating and
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC	The immunoglobulin component increases plasma half life and
CC	facilitates purification while deletion of the transmembrane region
CC	facilitates recovery, improves aqueous solubility and removes
CC	potentially immunogenic epitopes. Variants of the human lymphocyte
CC	cell surface glycoprotein described in AAP98106 are given in
CC	AAP98109-R98135. This variant contains a 241-Val-Glu-Asn insertion.
XX	
XX	
Sequence	374 AA;
50	



Query Match 100.0%; Score 215; DB 17; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGECEVETIINHTCNCVDGYGPGCQL 34  
|||||  
Db 159 CQWSCSGHGECEVETIINHTCNCVDGYGPGCQL 192

RESULT 33  
AAR98135  
ID AAR98135 standard; Protein; 374 AA.  
AC AAR98135;  
XX  
XX 01-NOV-1996 (first entry)  
DT  
XX  
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
DE  
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Region 1..37  
FT /label= Signal region.  
FT 38..154  
FT /label= Lectin domain.  
FT 159..192  
FT Domain /label= EGF domain.  
FT 196..257  
FT Binding-site /label= Complement binding repeat 1.  
FT 258..319  
FT Binding-site /label= Complement binding repeat 2.  
FT 335..357  
FT Domain /label= Transmembrane domain.  
FT 358..374  
FT /label= Cytoplasmic domain.  
XX  
XX US5514582-A.  
XX  
XX 07-MAY-1996.  
PD  
XX  
XX 23-FEB-1989; 89US-0315015.  
PE  
XX  
XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX  
XX Capon DJ, Lasky LA;  
PI  
XX  
XX WPI: 1996-238773/24.  
DR  
XX  
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
PT binding site of a receptor fused to Ig constant region - useful for  
PT diagnosis and treatment e.g. of inflammation  
FT  
XX  
XX Disclosure; Page 19; 41pp; English.  
PS  
XX  
XX A hybrid immunoglobulin chain comprising the ligand binding site of  
CC a single transmembrane receptor without an active transmembrane  
CC region; fused at its C-terminus with the N-terminus of an  
CC immunoglobulin constant region. The receptor is not a member of the  
CC immunoglobulin super family, nor a multiple subunit polypeptide  
CC encoded by discrete genes. The hybrid immunoglobulin chain combines

CC the adhesion/targeting of a ligand binding partner (LBP) with the  
CC effector functions of immunoglobulin and can bind to and/or activate  
CC more than one ligand. It can be used diagnostically for the in  
CC vitro assay of LBP and their targets, or therapeutically to deliver  
CC LBP such as toxins, enzymes, growth factors to particular cells.  
CC Typical applications are as antiviral, neuromodulating and  
CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
CC The immunoglobulin component increases plasma half life and  
CC facilitates purification while deletion of the transmembrane region  
CC facilitates recovery, improves aqueous solubility and removes  
CC potentially immunogenic epitopes. Variants of the human lymphocyte  
CC cell surface glycoprotein described in AAR98106 are given in  
CC AAR98109-R98135. This variant contains a 292-Tyr-Tyr-Tyr Insertion.  
XX  
XX  
SQ Sequence 374 AA;

Query Match 100.0%; Score 215; DB 17; Length 374;  
Best Local Similarity 100.0%; Pred. No. 3.1e-14;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 CQWSCSGHGECEVETIINHTCNCVDGYGPGCQL 34  
|||||  
Db 159 CQWSCSGHGECEVETIINHTCNCVDGYGPGCQL 192

RESULT 34  
AAR98115  
ID AAR98115 standard; Protein; 371 AA.  
AC AAR98115;  
XX  
XX 01-NOV-1996 (first entry)  
DT  
XX  
XX Human lymphocyte cell surface glycoprotein (HULHR) variant.  
DE  
XX Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation;  
KW metastasis.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Region 1..37  
FT /label= Signal region.  
FT 38..154  
FT /label= Lectin domain.  
FT 159..192  
FT Domain /label= EGF domain.  
FT 196..257  
FT Binding-site /label= Complement binding repeat 1.  
FT 258..316  
FT Binding-site /label= Complement binding repeat 2.  
FT 332..354  
FT Domain /label= Transmembrane domain.  
FT 355..371  
FT /label= Cytoplasmic domain.  
XX  
XX US5514582-A.  
XX  
XX 07-MAY-1996.  
PD  
XX  
XX 23-FEB-1989; 89US-0315015.  
PE  
XX  
XX 22-NOV-1989; 89US-0440625.  
PR 23-FEB-1989; 89US-0315015.  
PR 16-DEC-1991; 91US-0808122.  
PR 08-DEC-1992; 92US-0986931.  
PR 21-JAN-1994; 94US-0185670.  
XX  
XX (GETH ) GENENTECH INC.  
PA

```

XX  Capon DJ, Lasky LA;
PI
XX
XX  WPI, 1996-238773/24.
DR
XX
XX  Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT  binding site of a receptor fused to Ig constant region - useful for
PT  diagnosis and treatment e.g. of inflammation
XX
XX  Disclosure: Page 19; 41pp; English.
XX
XX  A hybrid immunoglobulin chain comprising the ligand binding site of
CC  a single transmembrane receptor without an active transmembrane
CC  region; fused at its C-terminus with the N-terminus of an
CC  immunoglobulin constant region. The receptor is not a member of the
CC  immunoglobulin super family, nor a multiple subunit polypeptide
CC  encoded by discrete genes. The hybrid immunoglobulin chain combines
CC  the adhesion/targeting of a ligand binding partner (LBP) with the
CC  effector functions of immunoglobulin and can bind to and/or activate
CC  more than one ligand. It can be used diagnostically for the in
CC  vitro assay of LBP and their targets; or therapeutically to deliver
CC  LBP such as toxins, enzymes, growth factors to particular cells.
CC  Typical applications are as antiviral, neuromodulating and
CC  immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC  treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC  The immunoglobulin component increases plasma half life and
CC  facilitates purification while deletion of the transmembrane region
CC  facilitates recovery, improves aqueous solubility and removes
CC  potentially immunogenic epitopes. Variants of the human lymphocyte
CC  cell surface glycoprotein described in AAR98106 are given in
CC  AAR98109-R98135. This variant contains an Ile174Leu substitution.
XX
XX  Sequence 371 AA;
SQ
XX
XX  Query Match 99.1%; Score 213; DB 17; Length 371;
XX  Best Local Similarity 97.1%; Pred. No. 4.8e-14;
XX  Matches 33; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 CQWSCSGHGECEVEIINNHTNCNDVGYGPOCOL 34
DB 159 cqpwschgceveelnhtncndvgygpcqql 192
XX
XX  RESULT 35
XX  AAR98116
XX  ID AAR98116 standard; Protein; 371 AA.
XX  AC AAR98116;
XX
XX  DT 01-NOV-1996 (first entry)
XX
XX  DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX  KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX  diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX  immunomodulator; cell adhesion; graft rejection; inflammation;
XX  metastasis.
XX
XX  OS Homo sapiens.
XX
XX  FH Key
XX  FT Region 1..37 Location/Qualifiers
XX  FT Domain /label= Signal region.
XX  FT Domain /label= Lectin domain.
XX  FT Domain /label= Lectin domain.
XX  FT Binding-site /label= EGF domain.
XX  FT Binding-site /label= Complement binding repeat 1.
XX  FT Binding-site /label= Complement binding repeat 2.
XX  FT Binding-site /label= Complement binding repeat 2.
XX  FT Domain 332..354

```

```

FT  /label= Transmembrane domain.
FT  Domain 355..371
FT  /label= Cytoplasmic domain.
XX
XX  US5514582-A.
XX
XX  07-MAY-1996.
XX
XX  23-FEB-1989; 89US-0315015.
XX
XX  22-NOV-1989; 89US-0440625.
XX  23-FEB-1989; 89US-0315015.
XX  16-DEC-1991; 91US-0808122.
XX  08-DEC-1992; 92US-0986931.
XX  21-JAN-1994; 94US-0185670.
XX
XX  (GETH ) GENENTECH INC.
XX
XX  Capon DJ, Lasky LA;
XX
XX  WPI, 1996-238773/24.
XX
XX  Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT  binding site of a receptor fused to Ig constant region - useful for
PT  diagnosis and treatment e.g. of inflammation
XX
XX  Disclosure: Page 19; 41pp; English.
XX
XX  A hybrid immunoglobulin chain comprising the ligand binding site of
CC  a single transmembrane receptor without an active transmembrane
CC  region; fused at its C-terminus with the N-terminus of an
CC  immunoglobulin constant region. The receptor is not a member of the
CC  immunoglobulin super family, nor a multiple subunit polypeptide
CC  encoded by discrete genes. The hybrid immunoglobulin chain combines
CC  the adhesion/targeting of a ligand binding partner (LBP) with the
CC  effector functions of immunoglobulin and can bind to and/or activate
CC  more than one ligand. It can be used diagnostically for the in
CC  vitro assay of LBP and their targets; or therapeutically to deliver
CC  LBP such as toxins, enzymes, growth factors to particular cells.
CC  Typical applications are as antiviral, neuromodulating and
CC  immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC  treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC  The immunoglobulin component increases plasma half life and
CC  facilitates purification while deletion of the transmembrane region
CC  facilitates recovery, improves aqueous solubility and removes
CC  potentially immunogenic epitopes. Variants of the human lymphocyte
CC  cell surface glycoprotein described in AAR98106 are given in
CC  AAR98109-R98135. This variant contains an Asn181Gln substitution.
XX
XX  Sequence 371 AA;
SQ
XX
XX  Query Match 97.2%; Score 209; DB 17; Length 371;
XX  Best Local Similarity 97.1%; Pred. No. 1.2e-13;
XX  Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 CQWSCSGHGECEVEIINNHTNCNDVGYGPOCOL 34
DB 159 cqpwschgceveelnhtncndvgygpcqql 192
XX
XX  RESULT 36
XX  AAR98114
XX  ID AAR98114 standard; Protein; 371 AA.
XX  AC AAR98114;
XX
XX  DT 01-NOV-1996 (first entry)
XX
XX  DE Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX
XX  KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX  diagnosis; therapy; drug delivery; antiviral; neuromodulator;

```

KW Immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..37  
 FT Domain /label= Signal region.  
 FT Domain 38..154  
 FT Domain /label= Lectin domain.  
 FT Domain 159..192  
 FT Domain /label= EGF domain.  
 FT Binding-site 196..257  
 FT Binding-site /label= Complement binding repeat 1.  
 FT Binding-site 258..316  
 FT Binding-site /label= Complement binding repeat 2.  
 FT Domain 332..354  
 FT Domain /label= Transmembrane domain.  
 FT Domain 355..371  
 FT Domain /label= Cytoplasmic domain.  
 XX  
 PN US5514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 XX  
 PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI: 1996-238773/24.  
 XX  
 XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation  
 XX  
 PS Disclosure; Page 19; 41pp; English.  
 XX  
 XX A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region. The receptor is not a member of the  
 CC immunoglobulin super family, nor a multiple subunit polypeptide  
 CC encoded by discrete genes. The hybrid immunoglobulin chain combines  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC treating graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC The immunoglobulin component increases plasma half life and  
 CC facilitates purification while deletion of the transmembrane region  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains an His168Gln substitution.  
 CC  
 XX Sequence 371 AA;  
 SQ

Query Match 96.3%; Score 207; DB 17; Length 371;  
 Best Local Similarity 97.1%; Pred. No. 1.9e-13;  
 Matches 33; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy 1 COPMSCSGHECEETIINNNHCNDGVYGGPOCOL 34  
 ||||||| ||||||| ||||||| ||||||| |||||||  
 Db 159 cqpwsctsgqgecevelnhntcndvgyy9pqcql 192

RESULT 37  
 AA57762  
 ID AA57762 standard; protein; 196 AA.  
 AC AA57762;  
 DT 17-MAR-2000 (first entry)  
 XX  
 DE Human L-selectin amino acid sequence.  
 XX  
 KW E-selectin; L-selectin; P-selectin; bifunctional molecule; regulation;  
 KW binding; detection; cellular proliferation; cytostatic; cell adhesion;  
 KW antiinflammatory; immunosuppressant; immunostimulator; cell migration;  
 KW tumour; neoplasia; cancer; inflammation.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 155  
 FT /note= "Unclear in specification"  
 XX  
 PN W09961033-A1.  
 XX  
 PD 02-DEC-1999.  
 XX  
 PF 21-MAY-1999; 99WO-US11300.  
 XX  
 PR 22-MAY-1998; 98US-0086442.  
 XX  
 PA (UYHO-) UNIV HOUSTON.  
 XX  
 PI Freidman J;  
 XX  
 DR WPI: 2000-072543/06.  
 XX  
 XX Novel bifunctional compounds, used for treating e.g. cancers,  
 PT inflammation or transplant rejection or for tissue repair or  
 PT regeneration -  
 XX  
 PS Example 2; Fig 3; 60pp; English.  
 XX  
 XX The present invention describes a bifunctional compound for the  
 CC regulation of cellular proliferation. The compound comprises, a  
 CC cell-adhesion oligosaccharide (CAOS), a linker group attached to the  
 CC CAOS by the reducing end of the CAOS or by a primary hydroxyl group,  
 CC and a nucleotide cyclic-3',5' monophosphate attached to the linker  
 CC group through a heterocyclic base. The bifunctional compounds can be  
 CC used for regulating cell proliferation, cell adhesion and cell  
 CC migration. Methods from the present invention can also be used to  
 CC identify agonists or antagonists of such activities. The bifunctional  
 CC compounds can be used for treating e.g. tumours, neoplasia, cancer or  
 CC inflammation. They can also be used to serve as immunosuppressants in  
 CC certain applications such as organ or tissue transplantation, as  
 CC immunostimulators and to provide immunological activity for disease  
 CC such as AIDS. They can also have important uses in the control and  
 CC generation and growth for new tissue. This can be useful in the repair  
 CC and regeneration of tissue as well as in the transplantation or  
 CC introduction of artificial tissue into the body. This can have  
 CC beneficial effects in the treatment of burns or other replacement  
 CC therapies requiring increased cell-cell adhesion, cell migration and/or  
 CC cellular growth. The present sequence represents a human L-selectin  
 CC amino acid sequence given in the exemplification of the present  
 CC invention.  
 XX  
 XX Sequence 196 AA;  
 SQ

Query Match 95.3%; Score 205; DB 21; Length 196;

Best Local Similarity 97.0%; Pred. No. 1.8e-13;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINHTCNCDDVGYGPOCQ 33  
Db 160 cqpwsqsgnhecveilnnyncncdvgygppcq 192

RESULT 38  
AAR22551

ID AAR22551 standard; Protein; 363 AA.

AC AAR22551;

DT 21-MAY-1992 (first entry)

DE T lymphocyte-specific Leu8 Antigen minor form.

KW Rapid immunoselection cloning technique; cell surface antigen;  
homolog receptor; antigen-presenting cells.

OS Homo sapiens.

FX Key Location/Qualifiers

FT Modified-site 113..115  
/label= N-linked\_glycosylation

FT Modified-site 157..159  
/label= N-linked\_glycosylation

FT Modified-site 230..232  
/label= N-linked\_glycosylation

FT Modified-site 269..271  
/label= N-linked\_glycosylation

FT Modified-site 285..287  
/label= N-linked\_glycosylation

FT Modified-site 299..301  
/label= N-linked\_glycosylation

FT Modified-site 324..326  
/label= N-linked\_glycosylation

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

FT Modified-site  
/note= "putative"

CC See AAQ21184 for the larger insert and AAR20815 for the major form of  
CC the Leu8 antigen that it encodes.  
XX

SQ Sequence 363 AA;

Query Match 95.3%; Score 205; DB 13; Length 363;  
Best Local Similarity 97.0%; Pred. No. 3e-13;  
Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 COPWSCSGHGECEIINHTCNCDDVGYGPOCQ 33  
Db 173 cqpwsqsgnhecveilnnyncncdvgygppcq 205

RESULT 39  
AAR91443

ID AAR91443 standard; Protein; 363 AA.

AC AAR91443;

DT 31-OCT-1996 (first entry)

DE Human Leu8 antigen.

KW Cell surface antigen; cloning; immunoselection; immunotherapy;  
therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;  
antiinflammatory.

OS Homo sapiens.

FX Key Location/Qualifiers

FT Modified-site 113..115  
/label= Glycosylation\_site

FT Modified-site 156..158  
/label= Glycosylation\_site

FT Modified-site 229..231  
/label= Glycosylation\_site

FT Modified-site 269..271  
/label= Glycosylation\_site

FT Modified-site 285..287  
/label= Glycosylation\_site

FT Modified-site 299..301  
/label= Glycosylation\_site

FT Modified-site 324..326  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

FT Modified-site  
/label= Glycosylation\_site

CC cloning method. The longer insert (AA114723) contained 2,350  
 CC residues, while the shorter lacked 436 internal residues. A major  
 CC transcript of 2.4 kb was present in peripheral blood mononuclear  
 CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,  
 CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukemic  
 CC T-cell lines. Leu8 antigens can be obtd. for diagnostic and  
 CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells  
 CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.  
 CC Soluble forms of Leu8 can act as antiinflammatory agents by reducing  
 CC lymphocyte migration.  
 CC  
 CC Sequence 363 AA;  
 SQ  
 Query Match 95.3%; Score 205; DB 17; Length 363;  
 Best Local Similarity 97.0%; Pred. No. 3e-13;  
 Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 COPWSCSGHGECEVEITNNHTCNCDDGYGPOCO 33  
 DB 173 cqpwsctghgeceveitnnhtcncddgygypqcc 205  
 RESULT 40  
 AA96184  
 ID AA96184 standard; Protein: 363 AA.  
 XX  
 AC AA96184;  
 XX  
 DT 19-DEC-2000 (first entry)  
 XX  
 DE Human T-cell specific Leu8 antigen.  
 XX  
 KW Leu8; cell surface antigen; human; immunoselection; panning;  
 KW immunodiagnosis; diagnosis; immunotherapy; gene therapy;  
 KW immune disorder; infection; asthma; immune-complex disease;  
 KW amyloidosis; multiple sclerosis; inflammation; antiinflammatory.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US611093-A.  
 XX  
 PD 29-AUG-2000.  
 XX  
 PE 28-OCT-1998; 98US-0181612.  
 XX  
 PR 01-DEC-1992; 92US-0983647.  
 PR 25-FEB-1988; 88US-0160416.  
 PR 13-JUL-1989; 89US-0379076.  
 PR 23-MAR-1990; 90US-0498809.  
 PR 13-JUL-1990; 90US-0535759.  
 XX  
 PA (GENO ) GEN HOSPITAL CORP.  
 XX  
 PT Stamenkovic I, Seed B;  
 XX  
 DR WPI: 2000-586382/55.  
 DR N-PSDB; AAA50632.  
 XX  
 PT Isolated nucleic acid molecule encoding the CD19 cell surface antigen,  
 PT useful for immunodiagnosis and immunotherapy of immune-mediated  
 PT infections or disorders, e.g. asthma, immune-complex disease, parasitic  
 PT diseases -  
 XX  
 PS Example 14; Column 69-72; 75pp; English.  
 XX  
 CC The present sequence is that of a T-cell specific Leu8 antigen, as  
 CC predicted from 1 of 2 clones (see AA50632) isolated from a human  
 CC T-cell library by the method of the invention. A longer Leu8  
 CC antigen is given in AA96184. The method, designed to isolate cell  
 CC surface antigen (CSA) cDNAs, is based upon transient expression of  
 CC a CSA in eukaryotic cells and physical selection of cells expressing  
 CC the antigen by adhesion to (panning on) an antibody-coated substrate

CC such as a culture dish. CSA nucleic acids isolated by the method of  
 CC the invention, and the proteins they encode, are useful for  
 CC immunodiagnosis and immunotherapeutic applications, including the  
 CC diagnosis and treatment of immune-mediated infections, diseases, and  
 CC disorders in animals, including humans. These disorders include  
 CC asthma, immune-complex disease, amyloidosis, parasitic diseases or  
 CC multiple sclerosis. The ability to interfere with the binding of  
 CC Leu8-T-cells to antigen presenting cells, or the ability to cause  
 CC such binding to occur on surfaces other than lymphocyte cells, can  
 CC be useful in diagnostics and therapy. The level of activated Leu8-  
 CC T-cells relative to resting Leu8+ cells could serve as a measure of  
 CC immune response to a particular antigen. Modification of the  
 CC specificity of the extracellular domain of Leu8, which mediates  
 CC adhesion to specific endothelial cells of lymph nodes, could serve  
 CC to regulate the homing potential of resting T cells. Soluble forms  
 CC of Leu8 could act as antiinflammatory agents by reducing lymphocyte  
 CC migration.  
 CC  
 CC Sequence 363 AA;  
 SQ  
 Query Match 95.3%; Score 205; DB 21; Length 363;  
 Best Local Similarity 97.0%; Pred. No. 3e-13;  
 Matches 32; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 COPWSCSGHGECEVEITNNHTCNCDDGYGPOCO 33  
 DB 173 cqpwsctghgeceveitnnhtcncddgygypqcc 205

Search completed: September 7, 2002, 10:14:07  
 Job time: 265 sec

**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:20 ; Search time 61.04 Seconds

(without alignments)  
374.104 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328

Perfect score: 733

Sequence: 1 CEPLAEDELGTMDCTHDFGN.....WSNPSPICQKLDKSFMIKE 132

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_19.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_ricent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	99.2	385	4 Q9UJ43	Q9UJ43 homo sapien
2	599	81.7	376	6 Q28629	Q28629 oryctolagus
3	535	73.0	372	11 Q63762	Q63762 rattus norv
4	318	43.4	646	6 Q29097	Q29097 sus scrofa
5	315	43.0	740	4 Q95508	Q95508 homo sapien
6	310	42.3	610	6 Q951G1	Q951G1 equus cabal
7	309	42.2	616	4 Q95509	Q95509 homo sapien
8	307	41.9	740	4 Q95507	Q95507 homo sapien
9	306	41.7	754	6 Q28290	Q28290 canis famill
10	300	40.9	609	6 Q9G1F0	Q9G1F0 canis famill
11	291	39.7	649	6 Q28657	Q28657 oryctolagus
12	290.5	39.6	485	6 Q951G3	Q951G3 odocoileus
13	275.5	37.6	484	6 Q951G2	Q951G2 sus scrofa
14	245	33.4	482	6 Q28982	Q28982 sus scrofa
15	192	26.2	3567	11 Q9ES77	Q9ES77 mus musculu
16	183	25.0	866	5 Q917E3	Q917E3 drosophila

17	183	25.0	933	5 Q917E4	Q917E4 drosophila
18	183	25.0	958	5 Q9V560	Q9V560 drosophila
19	173	23.6	1172	4 Q9NU87	Q9NU87 homo sapien
20	173	23.6	1653	5 Q9VITU9	Q9VITU9 drosophila
21	172.5	23.5	481	4 Q9H284	Q9H284 homo sapien
22	169	23.1	1236	11 Q91YB6	Q91YB6 rattus norv
23	168.5	23.0	1316	4 Q96J07	Q96J07 homo sapien
24	167.5	22.9	2014	6 Q29530	Q29530 pan troglod
25	163.5	22.3	2039	4 Q16745	Q16745 homo sapien
26	163.5	22.3	2489	4 Q16744	Q16744 homo sapien
27	160	21.8	3389	4 Q96C09	Q96C09 homo sapien
28	160	21.8	3508	4 Q96RM4	Q96RM4 homo sapien
29	156.5	21.4	390	11 Q921P0	Q921P0 mus musculu
30	155	21.1	533	11 Q08569	Q08569 cavia porce
31	155	21.1	555	11 Q99JY1	Q99JY1 cavia porce
32	154.5	21.1	222	6 Q19122	Q19122 callimico g
33	153	20.9	1911	6 Q29528	Q29528 papio cynoc
34	153	20.9	3564	11 Q923L3	Q923L3 mus musculu
35	152.5	20.8	363	6 Q02839	Q02839 sus scrofa
36	152.5	20.8	395	12 Q9J2M6	Q9J2M6 macaca mula
37	151.5	20.7	354	5 Q9U6I1	Q9U6I1 drosophila
38	151.5	20.7	410	4 Q9H4R2	Q9H4R2 homo sapien
39	151.5	20.7	465	4 Q60687	Q60687 homo sapien
40	150.5	20.5	560	5 Q22328	Q22328 caenorhabdl
41	150.5	20.5	1032	11 Q9DC83	Q9DC83 mus musculu
42	149.5	20.4	378	6 Q62837	Q62837 saquinus oe
43	149	20.3	315	6 Q28770	Q28770 papio cynoc
44	149	20.3	522	6 Q28769	Q28769 papio cynoc
45	148.5	20.3	559	4 Q9UQV2	Q9UQV2 homo sapien

## ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	385 AA.
1	Q9UJ43			
AC	Q9UJ43	01-MAY-2000 (TREMBLrel. 13, Created)		
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	L-SELECTIN PRECURSOR.			
GN	L-SELECTIN.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);			
RA	Fieger C.B.;			
RL	Submitted (Sep-1999) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=HEMATOPOIETIC (B LYMPHOCYTE);			
RA	Fieger C.B.;			
RL	Thesis (1998), Freie Universitaet Berlin, Fachbereich Chemie.			
DR	EMBL: AJ246000; CAB55488.1; -.			
DR	HSSP: P14151; IKDB.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001304; lectin_c.			
DR	InterPro: IPR002396; Selectin.			
DR	InterPro: IPR000436; Sushi_SCR_CCP.			
DR	Pfam: PF00008; EGF_1.			
DR	Pfam: PF00059; lectin_c_1.			
DR	Pfam: PF00084; sushi_1_2.			
DR	PRINTS: PR00343; SELECTIN.			
DR	SMART: SM00032; CCP_2.			
DR	SMART: SM00034; CLECT_1.			
DR	SMART: SM00181; EGF_1.			
DR	PROSITE: PS00615; C_TYPE_LECTIN_1; 1.			
DR	PROSITE: PS00041; C_TYPE_LECTIN_2; 1.			
DR	PROSITE: PS00022; EGF_1; UNKNOWN_1.			

DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.  
FT SIGNAL 1 51 POTENTIAL.  
FT CHAIN 52 385 L-SELECTIN  
SQ SEQUENCE 385 AA; 43617 MW; 1205f691BA638E91 CRC64;

Query Match 99.2%; Score 727; DB 4; Length 385;  
Best Local Similarity 99.2%; Pred. No. 2,1e-78;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60  
DB 210 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 269  
QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 120  
DB 270 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 329  
QY 121 QKDKSFSMIKE 132  
DB 330 QKDKSFSMIKE 341

RESULT 2  
ID 028629 PRELIMINARY; PRT; 376 AA.  
AC 028629;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE L-SELECTIN PRECURSOR.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_Taxid=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=KIDNEY CORTEX;  
RA Olan J., Marks R.M.;  
RT "cDNA for rabbit L-selectin.";  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U26535; AAA67896.1; -.  
DR HSSP; P14151; 1KJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin.c.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR00436; Sushl\_SCR\_CCP.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C-TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.  
FT SIGNAL 1 38 POTENTIAL.  
FT CHAIN 1 385 L-SELECTIN.  
SQ SEQUENCE 376 AA; 42346 MW; 59f6AD530F490947 CRC64;

Query Match 81.7%; Score 599; DB 6; Length 376;  
Best Local Similarity 80.3%; Pred. No. 3.8e-63;  
Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60  
DB 197 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 256

QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 120  
DB 257 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 316  
QY 121 QKDKSFSMIKE 132  
DB 317 QKDKSFSMIKE 328

RESULT 3  
ID 063762 PRELIMINARY; PRT; 372 AA.  
AC 063762;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE LYMFOCYTE MEMBRANE PROTEIN A.11.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_Taxid=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY;  
RX MEDLINE=95369821; PubMed=7543874;  
RA Sackstein R., Meng L., Xu X.M., Chin Y.H.;  
RT "Evidence of post-transcriptional regulation of L-selectin gene  
RT expression in rat lymphoid cells.";  
RL Immunology 85:198-204(1995).  
DR EMBL; S79523; AAC60710.2; -.  
DR HSSP; P14151; 1KJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin.c.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR00436; Sushl\_SCR\_CCP.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 2.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C-TYPE\_LLECTIN\_1; 1.  
DR PROSITE; PS50041; C-TYPE\_LLECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein.  
SQ SEQUENCE 372 AA; 42471 MW; 7DFD125610DD6E4A CRC64;

Query Match 73.0%; Score 535; DB 11; Length 372;  
Best Local Similarity 71.2%; Pred. No. 1.6e-55;  
Matches 94; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 60  
DB 197 CEPLAPDLGMDCTHPGNFSSQCAFSCSEGTNLGIEETTCGPRGNMSSPEPTCOV 256  
QY 61 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 120  
DB 257 IQCEPLAPDGLGIMNCSPHLASFSTACTFCISGTELLGKKTICSSGIMWSPSPIC 316  
QY 121 QKDKSFSMIKE 132  
DB 317 QKDKSFSMIKE 328

RESULT 4  
ID 029097 PRELIMINARY; PRT; 646 AA.  
AC 029097;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)



DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE P-SELECTIN PRECURSOR.  
 OS Sus scrofa (pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 NC NCBITaxid=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-AORTA.  
 RA Rollins S.A., Johnson K.K., Birks C.W., Matis L.A., Rother R.P.;  
 RL Submitted (Oct-1995) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-AORTA;  
 RX MEDLINE=20171534; PubMed=10706724;  
 RA Stocker C.J., Sugars K.L., Harari O.A., Landis R.C., Morley B.J.,  
 RA Haskard D.O.;  
 RT "TNF-alpha, IL-4, and IFN-gamma regulate differential expression of P-  
 RT and E-selectin expression by porcine aortic endothelial cells";  
 RL J. Immunol. 164:3309-3315(2000).  
 DR EMBL: L39075; AAU79007.1;  
 DR EMBL: AF163766; AAU43272.1;  
 DR HSSP: P16109; IFSB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001304; lectin\_c.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 6.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 6.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00615; C\_TYPE\_LECTIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS00041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW EGF-like domain; Glycoprotein; Lectin; Selectin; Signal.  
 FT SIGNAL  
 FT CHAIN 1 41 POTENTIAL.  
 FT SEQUENCE 646 AA; 71127 MW; 3863FAFE09F0BB6 CRC64;

Query Match 43.4%; Score 318; DB 6; Length 646;  
 Best Local Similarity 42.1%; Pred. No. 2.2e-29;  
 Matches 51; Conservative 22; Mismatches 48; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPNFSFSSQAFSCSEGTNLGTIEFTGCPFGNWSPEPTCOV 60  
 DB 324 CEPLSPVAGSMDCPFSSSAFOYNTSCSRCAKGLTRADYVRCNSLQGTAPAPVCA 383  
 QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFTCSGTELGKTKTICSSGIMSNSPIC 120  
 DB 384 LQCODLPAPKAOVNCVSHFAGFRVQSTGCTDCGSSLVGASVLCLEFGTGMASAPAPC 443  
 QY 121 Q 121  
 DB 444 Q 444  
 RESULT 5  
 095508 PRELIMINARY: PRT; 740 AA.  
 AC 095508;  
 DT 01-MAY-1999 (TREMBlrel. 10, Created)  
 DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE DJ78OM13.1.2 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN  
 DE CD62, GMP140)) (ISOFORM 2) (FRAGMENT).  
 GN SELP.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBITaxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Howden P.;  
 RL Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AL022146; CAU18143.1;  
 DR HSSP: P16109; IKUD.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001304; lectin\_c.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00008; EGF; 1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 8.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 8.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00615; C\_TYPE\_LECTIN\_1; 1.  
 DR PROSITE: PS00041; C\_TYPE\_LECTIN\_2; 1.  
 DR PROSITE: PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 KW EGF-like domain; Glycoprotein.  
 FT NON\_TER 740  
 FT SEQUENCE 740 AA; 81390 MW; 1D2E35E6D93745CE CRC64;

Query Match 43.0%; Score 315; DB 4; Length 740;  
 Best Local Similarity 43.9%; Pred. No. 6e-29;  
 Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPNFSFSSQAFSCSEGTNLGTIEFTGCPFGNWSPEPTCOV 60  
 DB 262 COHLEAPSEGTMDCTHPNFSFSSQAFSCSEGTNLGTIEFTGCPFGNWSPEPTCOV 321  
 QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFTCSGTELGKTKTICSSGIMSNSPIC 120  
 DB 322 IOCEPLSPVAGSMDCPFSSSAFOYNTSCSRCAKGLTRADYVRCNSLQGTAPAPVCA 381  
 QY 121 QRL 123  
 DB 382 QAL 384

RESULT 6  
 095LGI PRELIMINARY: PRT; 610 AA.  
 AC 095LGI;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Equus caballus (Horse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.  
 NC NCBITaxid=9796;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=2142134; PubMed=11529941;  
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
 RA MacLachlan N.J.;  
 RT "Characterization of equine E-selectin";  
 RL Immunology 103:498-504(2001).  
 DR EMBL: AF307972; AAK48712.1;  
 KW Lectin; Selectin.  
 FT SEQUENCE 610 AA; 66191 MW; F9D3DD12C445382 CRC64;

Query Match 42.3%; Score 310; DB 6; Length 610;  
 Best Local Similarity 44.6%; Pred. No. 1.9e-28;  
 Matches 54; Conservative 16; Mismatches 51; Indels 0; Gaps 0;

OY		1	CPELPAEELCTMDCDTHFGNPFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV	60
		:	:             :     :	
Db		179	CQAGAEAFEHRLVCTHFLGNFYSVSSCSVEEYLPSKREAMOCSTSGENASAPPACHV	238
OY		61	IQCCEPLSAPDLGIMNCSHPLASFSTACTFTCSGETELIGKKTKICSSGIWNSPAPIC	120
		:	:  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :	
Db		239	VECDALTNPNANGVMWCOSQSDFPMNTTCTFDCCGEFELNGPHLAQTCPSONMDNEKPTC	298
OY		121	Q 121	
		:		
Db		299	K 299	
RESULT		7		
ID	095509		PRELIMINARY;	PRT; 616 AA.
AC	095509;			
DT	01-MAY-1999 (TREMBLrel, 10, Created)			
DT	01-MAY-1999 (TREMBLrel, 10, last sequence update)			
DT	01-DEC-2001 (TREMBLrel, 19, last annotation update)			
DE	DJ780ML3.1.3 (SELECTIN P (GRANULE MEMBRANE PROTEIN 140 KD, ANTIGEN CD62, GMP140)) (ISOFORM 3) (FRAGMENT).			
GN	SELF.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Howden P;			
RL	Submitted (NOV-1998) to the EMBL/GenBank/DDBJ databases.			
DR	EMBL; AL022146; CAA18144.1; -			
DR	HSP; P16109; IKJD.			
DR	InterPro: IPR000561; EGF-like.			
DR	InterPro: IPR001304; lectin.c.			
DR	InterPro: IPR002396; Selectin.			
DR	InterPro: IPR000436; Sushi_SCR_CCP.			
DR	Pfam; PF00008; EGF_1.			
DR	Pfam; PF00059; lectin_c; 1.			
DR	Pfam; PF00084; sushi; 6.			
DR	PRINTS; PR00343; SELECTIN.			
DR	SMART; SM00032; CCP; 6.			
DR	SMART; SM00034; CLECT; 1.			
DR	PROSITE; PS00615; C_Type_Lectin_1; 1.			
DR	PROSITE; PS50041; C_Type_Lectin_2; 1.			
DR	PROSITE; PS00022; EGF_1; UNKNOWN_1.			
DR	PROSITE; PS01186; EGF_2; 1.			
DR	EGF-like domain; Glycoprotein.			
KW	NON TER			
FT	616			
SQ	SEQUENCE 616 AA; 67736 MW; 35CDABFADE6ID72A CRC64;			
Query Match		42.2%; Score 309; DB 4; Length 616;		
Best Local Similarity		41.5%; Pred. No. 2.5e-28;		
Matches	51; Conservative	23; Mismatches 49; Indels 0; Gaps 0;		
OY		1	CPELPAEELGTMDCHTPRGNFSFSQCAFCSEGTNLGIEETTCGPGNMSSPEPTCOV	60
		:	:  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :	
Db		200	GGELELPCHVLNMCNCHPIGNFSFNQCSFHCTDGYQVGNPKSLCLELAGIWTNKPPQCIA	259
OY		61	IQCCEPLSAPDLGIMNCSHPLASFSTACTFTCSGETELIGKKTKICSSGIWNSPAPIC	120
		:	:  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :  : :	
Db		260	AQCPEPLKIPERGNMTCLISAKAFHQHSQCSFSCDEGFAIVGVQCTASGWTAAPVC	319
OY		121	QKL 123	
		:		
Db		320	KAL 322	
RESULT		8		
ID	095507		PRELIMINARY;	PRT; 740 AA.

AC 095507;  
DT 01-MAY-1999 (TREMBLrel. 10, Created)  
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE D7J80M3.1.1 (SELECTIN\_P, GRAVULE MEMBRANE PROTEIN 140 KD, ANTIGEN  
DE CD62, GMP140) (ISOFORM 1) (FRAGMENT).  
SELP.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Howden P.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL022146; CAA18142.1; -  
DR HSSP; P16109; 1KTD.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR001304; lectin\_c.  
DR InterPro; IPR002396; Selectin.  
DR InterPro; IPR000436; Sush1\_SCR\_CCP.  
DR Pfam; PF00008; EGF\_1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sush1; 8.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 8.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF\_1.  
DR PROSITE; PS00615; C\_TYPE\_LECTIN\_1; 1.  
DR PROSITE; PS50041; C\_TYPE\_LECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
FT EGF-like domain; Glycoprotein.  
KW NON\_TER 740  
KW 740  
SQ SEQUENCE 740 AA; 81095 MW; 3B5F70A45B1A3CD4 CRC64;

Query Match	41.9%	Score 307	DB 4	Length 740
Best Local Similarity	40.7%	Pred. No. 5,4e-28		
Matches 50	Conservative 24	Mismatches 49	Indels 0	Gaps 0
Qy 1	CEPLEAELETMTDCTHPFNFSFSSQCAFCSEGTNLGTGIEETTCGPGFMWSSPEPTQY 60	: : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :		
Dy 200	CGEELPQPHYHLMNCSPHLNFSFNQCSCHCTDGTQVNGVSKLECLASGIWTKKPPQCIA 259	: : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :		
Qy 61	IOCEPLASAPDLGTMCNSHPLASFSFTSACTFCISGTELLIGKKRTTICESSGIWSSNPIC 120	: : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :		
Dy 260	AOCGPPLKIPERGMNMTCLHSAKAFHOSSCSFSGEGFALVPEVYQCTAGVWTAPAPVC 319	: : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :   : : :		
Qy 121	OKL 123	: :		
Dy 320	KAI 322	:		
RESULT 9				
Q28290				
ID Q28290	PRELIMINARY	PRT	754	AA.
AC Q28290				
DT 01-NOV-1996	(TREMBLrel. 01, Created)			
DT 01-NOV-1996	(TREMBLrel. 01, last sequence update)			
DT 01-DEC-2001	(TREMBLrel. 19, last annotation update)			
DE CEL1 ADHESION MOLECULE PRECURSOR (FRAGMENT).				
GN GMP140.				
OS Canis familiaris (Dog).				
OC Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;				
OC Mammalia; Eutheraia; Carnivora; Fissipedia; Canidae; Canis.				
OC NCBI_TaxID=9615;				
OX [1]				
RA BESEQUENCE FROM N.A.				
RC TISSUE=SPLEEN.				
RA Manning A.M., Sanders W.E.Jr., Kukielka G.L., Dore M.,				
RA Rosenbloom C.L., Hawkins H.L., Michael L.H., Entman M.L., Smith C.W.,				
RA Beaudet A.L., Anderson D.C.;				

"Molecular cloning of canine GMP140 and studies of expression in a model of myocardial ischemia/reperfusion.";  
RT Submitted (MAR-1995) to the EMBL/Genbank/DBJ databases.  
RL EMBL; M88170; AAA63789.1; -.  
DR HSSP; P16109; IFSB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin\_C.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR Pfam; PF00084; sush1; 8.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 8.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C-type-LECTIN\_1; 1.  
DR PROSITE; PS50041; C-type-LECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein; Signal.  
FT NON\_TER 1  
FT SIGNAL <1 13 POTENTIAL.  
FT CHAIN 14 754 CELL ADHESION MOLECULE.  
SQ SEQUENCE 754 AA; 82303 MM; F0438BEAA521E773 CRC64;  
  
Query Match 41.7%; Score 306; DB 6; Length 754;  
Best Local Similarity 40.5%; Pred. No. 7.3e-28;  
Matches 49; Conservative 24; Mismatches 48; Indels 0; Gaps 0;  
  
Qy 1 CEPLAPELGTMDCTHPGNFSSSOCAFCSEGTNLGIEETTCGPGMSSPPTCOV 60  
Db 185 CGDFPLPHVLMNCNHPJGNSFSNCSFYCTEGTELGPSPLECLASTMTNKPRCVA 244  
Qy 61 IQCEPLASPDIGIMNCSPHLASFSTSACTFCISGTELLIGKKTKICSSGIGMNSPPTC 120  
Db 245 IQCEPLKTPREGSMNCLHVSFAFYQSSCHFCSEGFALYGEVYQCTASGMWTAAPVC 304  
Qy 121 Q 121  
Db 305 E 305  
  
RESULT 10  
Q9GLF0 PRELIMINARY; PRT; 609 AA.  
AC Q9GLF0;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE E-SELECTIN.  
OS Canis familiaris (Dog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.  
OX NCBI\_TaxID=9615;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-ENDOTHELIAL CELLS;  
RA Zheng L., Shi Y., Wu H., Zhang G.;  
RT "Cloning and sequencing of beagle E-selectin genomic DNA and  
RT comparison with other species.";  
RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF287257; AAG10039.1; -.  
DR HSSP; P16581; IFSB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin\_C.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR Pfam; PF00059; lectin\_C; 1.  
DR Pfam; PF00084; sush1; 6.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 6.  
DR SMART; SM00034; CLECT; 1.

DR SMART; SM00181; EGF; 4.  
DR SMART; SM00001; EGF-like; 1.  
DR PROSITE; PS50041; C-type-LECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
KW lectin; Selectin.  
SQ SEQUENCE 609 AA; 66073 MM; 41E62D1FAD23881F CRC64;  
  
Query Match 40.9%; Score 300; DB 6; Length 609;  
Best Local Similarity 43.0%; Pred. No. 2.9e-27;  
Matches 52; Conservative 18; Mismatches 51; Indels 0; Gaps 0;  
  
Qy 1 CEPLAPELGTMDCTHPGNFSSSOCAFCSEGTNLGIEETTCGPGMSSPPTCOV 60  
Db 181 COAEPPEHGSIVCTHPLCTFSYNSCFYCKGILPSTETVQCTSTGEMSPACNV 240  
Qy 61 IQCEPLASPDIGIMNCSPHLASFSTSACTFCISGTELLIGKKTKICSSGIGMNSPPTC 120  
Db 241 VECNALTNPCNCHGVMDLQSSGNFPNNMTCFECCEGFELMGPKRLQCTSSGMWDRKRPIC 300  
Qy 121 Q 121  
Db 301 K 301  
  
RESULT 11  
Q28657 PRELIMINARY; PRT; 649 AA.  
AC Q28657;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE P-SELECTIN.  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Vora D.K., Fang Z., Liya S.M., Parham F., Watson A.D., Drake T.A.,  
RA Territo M.C., Berliner J.A.;  
RT "Induction of P-selectin by KM-LDL and its role in human  
RT atherosclerosis.";  
RL Submitted (NOV-1995) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Warden C.H.;  
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.  
DR EMBL; U39446; AAB81385.1; -.  
DR HSSP; P16109; IFSB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR001304; lectin\_C.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR Pfam; PF00059; EGF; 1.  
DR Pfam; PF00084; sush1; 6.  
DR Pfam; PF00084; sush1; 6.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 6.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00615; C-type-LECTIN\_1; 1.  
DR PROSITE; PS50041; C-type-LECTIN\_2; 1.  
DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
DR PROSITE; PS01186; EGF\_2; 1.  
KW EGF-like domain; Glycoprotein; Lectin; Selectin.  
SQ SEQUENCE 649 AA; 71755 MM; ECCD8C847B84BC31 CRC64;  
  
Query Match 39.7%; Score 291; DB 6; Length 649;  
Best Local Similarity 40.0%; Pred. No. 3.7e-26;  
Matches 48; Conservative 25; Mismatches 47; Indels 0; Gaps 0;

QY 4 LEAPLGTMDCTHPGNSFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 63  
 Db 203 LDLPQHVMNCSHPLGNFSFNHSCSFHCADYALNGPSELECLASGIMTNSPQCVAAQC 262  
 QY 64 EPLAPDGLGIMNCNCHPLASFSFTSACTFTCSGTCLIGKKTTCSSSISWNSPSPICOL 123  
 Db 263 PALKSPEGSMSCVQASAEAFHQSSCSFSCBEGFELVGEVYHCTALGVMTAPFPVCKAL 322

## RESULT 12

Q95LG3 PRELIMINARY; PRT; 485 AA.  
 AC Q95LG3;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Odocoileus hemionus (Mule deer) (Black-tailed deer).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Cervidae;  
 OC Cervidae; Odocoileinae; Odocoileus.  
 OX NCBI\_TaxID=9872;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21421234; PubMed=11529941;  
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
 RT MacLachlan N.J.;  
 RT "Characterization of equine E-selectin."  
 RL EMBL; AF307970; AAK48710.1; -.  
 DR EMBL; AF307970; AAK48710.1; -.  
 KW Lectin; Selectin.  
 SQ SEQUENCE 485 AA; 53247 MW; 69959199EAFE9980 CRC64;

Query Match 39.6%; Score 290.5; DB 6; Length 485;  
 Best Local Similarity 43.8%; Pred. No. 3.1e-26;

Matches 53; Conservative 19; Mismatches 48; Indels 1; Gaps 1;

QY 1 CEPLAPELLGTMDCTHPGNSFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 60  
 Db 181 CQAKHPEHGHVLC-NPLGKFTYNSCSISCAEGYLPSTEARCKMSSGEMSTPLPCNV 239  
 QY 61 IQCEPLAPDGLGIMNCNCHPLASFSFTSACTFTCSGTCLIGKKTTCSSSISWNSPSPIC 120  
 Db 240 VKCDALSNPDNGVYVNCNSQNGSLPWNMTCTFECEGKYLGPQHLQCTSSGIMDNKQPTC 299  
 QY 121 Q 121  
 Db 300 K 300

## RESULT 13

Q95LG2 PRELIMINARY; PRT; 484 AA.  
 AC Q95LG2;  
 DT 01-DEC-2001 (Tremblrel. 19, Created)  
 DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21421234; PubMed=11529941;  
 RA Hedges J.F., Demaula C.D., Moore B.D., McLaughlin B.E., Simon S.I.,  
 RT MacLachlan N.J.;  
 RT "Characterization of equine E-selectin."  
 RL EMBL; AF307971; AAK48711.1; -.  
 DR EMBL; AF307971; AAK48711.1; -.  
 KW Lectin; Selectin.

SQ SEQUENCE 484 AA; 52951 MW; 617930C1C2F47B44 CRC64;

Query Match 37.6%; Score 275.5; DB 6; Length 484;  
 Best Local Similarity 40.5%; Pred. No. 1.9e-24;

Matches 49; Conservative 21; Mismatches 50; Indels 1; Gaps 1;

QY 1 CEPLAPELLGTMDCTHPGNSFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 60  
 Db 180 CQAKHPEHGHVLC-NPLGKFTYNSCSISCAEGYLPSTEARCKMSSGEMSTPLPCNV 238  
 QY 61 IQCEPLAPDGLGIMNCNCHPLASFSFTSACTFTCSGTCLIGKKTTCSSSISWNSPSPIC 120  
 Db 239 VKCDALSNPDNGVYVNCNSQNGSLPWNMTCTFECEGKYLGPQHLQCTSSGIMDNKQPTC 298  
 QY 121 Q 121  
 Db 299 K 299

## RESULT 14

Q28982 PRELIMINARY; PRT; 482 AA.  
 AC Q28982;  
 DT 01-NOV-1996 (Tremblrel. 01, Created)  
 DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)  
 DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE E-SELECTIN.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97075911; PubMed=8918234;  
 RA Winkler H., Brostjan C., Csizmadia V., Natarajan G., Anrather J.,  
 RA Bach F.H.;  
 RT "The intron-exon structure of the porcine E-selectin-encoding gene."  
 RL Gene 176:67-72(1996).  
 DR EMBL; U37521; AAC48680.1; -.  
 DR HSSP; P16581; IESL.  
 DR InterPro; IPR000561; EGF-like.  
 DR InterPro; IPR001304; Lectin\_C.  
 DR InterPro; IPR002396; Selectin.  
 DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam; PR00008; EGF\_1.  
 DR Pfam; PR00059; Lectin\_C; 1.  
 DR Pfam; PR00084; sushi; 4.  
 DR PRINTS; PR00343; SELECTIN.  
 DR SMART; SM00032; CCP; 4.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
 DR PROSITE; PS50041; C-TYPE\_LECTIN\_2; 1.  
 DR PROSITE; PS00022; EGF\_1; UNKNOWN\_1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 KW EGF-like domain; Glycoprotein; Lectin; Selectin.  
 SQ SEQUENCE 482 AA; 52341 MW; 97DC5D70B8115944 CRC64;

Query Match 33.4%; Score 245; DB 6; Length 482;  
 Best Local Similarity 35.0%; Pred. No. 8.1e-21;

Matches 43; Conservative 22; Mismatches 56; Indels 2; Gaps 2;

QY 1 CEPLAPELLGTMDCTHPGNSFSSQCAFSCSEGTNLGTIEETTCGPPGNMSSPEPTCOV 59  
 Db 240 CDVGHHPQNGDVSNCHSSISGEFAVYKSTCHFTCAEGFGLOGPAOICTAOGQWTOQAPVCK 299  
 QY 60 VIQCEPLAPDGLGIMNCNCHPLASFSFTSACTFTCSGTCLIGKKTTCSSSISWNSPSPIC 118  
 Db 300 AKKCAVSPQKNGLVKFTHTSPGTGFTYVSSCAFSCEBGEELKGSQQLACTSGOGWTOEVP 359  
 QY 119 IQ 121

```
DB 360 SCQ 362

RESULT 15
ID Q9ES77 PRELIMINARY; PRT; 3567 AA.
AC Q9ES77;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE POLYDOM PROTEIN PRECURSOR.
GN POLYDOM.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C3H/HENSIC;
RX MEDLIN=20517255; PubMed=11062057;
RA Gilles D., Vinit M.-A., Callebaut I., Coulombel L., Cacheux V.,
RA Romeo P.-H., Vigon I.;
RT "Polydom : a secreted protein with pentraxin, complement control
RT protein, epidermal growth factor and von willebrand factor A
RT domains.";
RL Blochem. J. 352:49-59(2000).
DR EMBL: AF206329; AAC32160.1; -.
DR HSSP: P00740; IEDM.
DR MGD: MGI:1928849; Polydom.
DR InterPro: IPR000152; Asx_hydroxyl.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR00742; EGF_2.
DR InterPro: IPR01881; EGF_Ca.
DR InterPro: IPR001438; EGF-II.
DR InterPro: IPR003410; HYR.
DR InterPro: IPR001759; Pentaxin.
DR InterPro: IPR000436; Sushl_SCR_CCP.
DR InterPro: IPR002035; VWFA.
DR Pfam: PF00008; EGF; 10.
DR Pfam: PF02494; HYR; 2.
DR Pfam: PF00084; sushl; 33.
DR Pfam: PF00092; vwa; 1.
DR PRINTS: PR00010; EGFBLD.
DR PRINTS: PR00895; PENTAXIN.
DR PRINTS: PR00453; VWFADOMAIN.
DR PRODOM: PD002153; Pentaxin; 1.
DR SMART: SM00032; CCP; 34.
DR SMART: SM00181; EGF_CA; 9.
DR SMART: SM00179; EGF_CA; 15.
DR SMART: SM00159; PTH; 1.
DR SMART: SM00327; VWFA; 1.
DR PROSITE: PS00010; ASX_HYDROXYL; UNKNOWN_1.
DR PROSITE: PS50234; VWFA; 1.
KW Signal.
FT SIGNAL.
SO SEQUENCE 3567 AA; 387391 MW; 8FBA8276E12293E5 CRC64;

Query Match 26.2%; Score 192; DB 11; Length 3567;
Best Local Similarity 27.1%; Pred. No. 1.7e-13;
Matches 49; Conservative 17; Mismatches 51; Indels 64; Gaps 5;

QY 1 CEPLEAPELGTMCTHPFG-NFSSSCAFSCSEGTNLGTIEETTCGPFGNMSSEPTCQ 59
DB 2140 CIPVRCGEPPSTANGVYSGSTVNSFGAVVAYSCHKGFYIKGKSTCEAAGQSKPTPTCH 2199
QY 60 VIQCE-----PLSAP 69
DB 2200 PVSCKNEPPKVENGFLEHTTGTTFESEARQCNPQYKAAGSPVYVCOAHNMHSDAPLSCT 2259
QY 70 DLGINCSHPL-----ASFSTSACTFCISGTETLICKKTKICSSGIWS-NPSPI 119
```

```
DB 2260 PL---NCGKPPPTIONGFLKGSFEVGSKKQGFVYCNNGYELVGNMSWTCKSKGSKMPPSK 2316
QY 120 C 120
DB 2317 C 2317

RESULT 16
ID Q917E3 PRELIMINARY; PRT; 866 AA.
AC Q917E3;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HIG PROTEIN.
GN HIG OR CG2040.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLIN=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abail J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Botkova D.C., Buchan M.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Butts J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Hejman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
RA Palazzolo M., Plattman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan K.C., Zhan M.G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003834; AAC22297.1; -.
DR HSSP: P10998; IYD.
DR FlyBase: FBgn0010114; hlg.
DR InterPro: IPR003599; Ig.
DR InterPro: IPR003006; Ig_MHC.
DR InterPro: IPR002396; Selection.
DR InterPro: IPR000436; Sushl_SCR_CCP.
DR Pfam: PF00047; Ig; 1.
DR Pfam: PF00084; sushl; 4.
```



	Pfam:	PF00084;	sush1	: 19.
DR	SMART:	SM00032;	CCP	: 19.
DR	SEQUENCE	1172 AA;	132087 MW;	8F5B954C4BAFA454 CRC64;

  

Query Match	23.6%;	Score 173;	DB 4;	Length 1172;
Best Local Similarity	28.8%;	Pred. No. 8,7e-12;		
Matches 36;	Conservative 25;	Mismatches 48;	Indels 16;	Gaps 5;

  

Dy	14 CTH----	PGNFS-----	FSSQCAFSCSEGTNLTG--IEYTCGPRGNMSSPEPTCOV	60
	:	:	:	:
Dd	85 CGHPEDTFFGGFTLT	TGCNVAFEGYKAVYTCKNEGYYLLTEINVR	ECDTDG--WTNDIPICEV	143
	:	:	:	:
Dy	61 IQCEPLSPADLG--	IIMNSHPLASSFPISACTFCITCTELIGKKKTIC	CESSGIWMSNP	118
	:: ::     :   :	:   :   :   :	:   :   :   :	:
Dd	144 VKCLEVTATPENGLIVS	SAMEPDREIHFQAAVFNCNSGYKITGEDMHCSDDGEFWSKEP		203
	:	:	:	:
Dy	119 ICOKL 123			
	:			
Dd	204 KCVEI 208			

  

RESULT	20
O9VIU9	
ID	O9VIU9 PRELIMINARY; PRT; 1653 AA.
DC	O9VIU9:
DT	01-MAY-2000 (TREMBLE) 13 Created)
DT	01-MAR-2001 (TREMBLE) 16 Last sequence update)
DT	01-DEC-2001 (TREMBLE) 19 Last annotation update)
DE	CG10186 PROTEIN.
DN	CG10186.
OS	Drosophila melanogaster (fruit fly).
OC	Eukaryota Metazoa Arthropoda Tracheata Hexapoda Insecta;
OC	Pterygota Neoptera Endopterygota Diptera Brachycera Muscomorpha;

[1]

RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
Ra George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
Ra George G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
Ra Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
Ra Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
Ra Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
Ra Baller R.M., Bauman A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
Ra Beeson K.Y., Bernos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
Ra Borokov D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
Ra Butts K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
Ra Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
Ra de Pallos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
Ra Dodson K., Doup L.E., Downes M., Dugan-Hochia S., Dunkov B.C., Dunn P.,  
Ra Dudin K.J., Evangelista C.C., Ferriz C., Ferreira S., Fleischmann W.,  
Ra Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
Ra Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
Ra Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
Ra Hostin D., Houston K.A., Howland T.J., Wei M.-H., Iobagyan C.,  
Ra Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemnison J.A., Ketchum K.A.,  
Ra Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
Ra Liao X., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
Ra Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,  
Ra Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
Ra Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
Ra Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pecled J.M.,  
Ra Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
Ra Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
Ra Shue B.C., Siden-Kiamos I., Sampson M., Skupski M.P., Smith T.,  
Ra Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
Ra Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
Ra Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissensebach J.,  
Ra Williams S.M., Woodage T., Wooley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.W., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glibos R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AF003663; AAF53815.2; -  
 DR HSSP: 046655; ICBH.  
 DR FLYBase: FBgn0032797; CG10186.  
 DR InterPro: IPR000923; Copper\_blue1.  
 DR InterPro: IPR000566; Lipoclin\_cyFABP.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR002221; WAP.  
 DR Pfam: PF00084; sushi; 14.  
 DR Pfam: PF00095; wap; 1.  
 DR SMART: SM00032; CCP; 17.  
 DR SMART: SM00217; WAP; 1.  
 DR PROSITE: PS00196; COPPER\_BLUE; UNKNOWN\_1.  
 DR PROSITE: PS00213; LIPOCALIN; UNKNOWN\_1.  
 SQ SEQUENCE 1653 AA; 180864 MW; 174E7775C51FD60F CRC64;

Query Match 23.6%; Score 173; DB 5; Length 1653;  
 Best Local Similarity 25.4%; Pred. No. 1.3e-11;  
 Matches 48; Conservative 16; Mismatches 49; Indels 76; Gaps 5;

QY 1 CEPLAELGTMDCTHHPGNFSFSS-----QCAFSCSEGTNLGIEETTCGP 47  
 DB 1437 CESVEECGDI-----PLGMSNASPRVSVLSREVGRAFAFSCAGYGLGPAALICNP 1489  
 QY 48 FGNSSPEPTCOVIOCEPLSAPDGL----- 72  
 DB 1490 TGEWSAPLPTCVCEVQCPNGAPQNGVAGSAPYAGDVQFNPCEYMGQPIIACODN 1549  
 QY 73 -----IMNCSDP-----LASFSTSA---CTFICSGTELIGKKTICSSG 111  
 DB 1550 ARMSGGLPKVCVQACSYPGTVISGRMSSVKFYAIGESTITFTCDAGLDLRGSKVLAKLKG 1609  
 QY 112 IWSNPSPTC 120  
 DB 1610 KWSAIPTC 1618

RESULT 21  
 Q9H284 PRELIMINARY; PRT; 481 AA.  
 AC 09H284;  
 DT 01-MAR-2001 (TREMblrel. 16, Created)  
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE SEROLOGICALLY DEFINED BREAST CANCER ANTIGEN NY-BR-38.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BREAST;  
 RA Seanlan M.J., Gout I., Stockert E., Gure A.O., Jaeger D., Chen Y.-T.,  
 RA Old L.J.;  
 RT "Humoral Immunity to Human Breast Cancer: Antigen Definition and  
 RT Quantitative Analysis of mRNA Expression.";  
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF308289; AAG48257.1; -  
 DR HSSP: P10998; IYVD.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR Pfam: PF00084; sushi; 7.  
 DR SMART: SM00032; CCP; 8.  
 SQ SEQUENCE 481 AA; 51891 MW; FD99724DEF7C69B0B CRC64;

Query Match 23.5%; Score 172.5; DB 4; Length 481;  
 Best Local Similarity 31.2%; Pred. No. 3.6e-12;  
 Matches 39; Conservative 18; Mismatches 59; Indels 9; Gaps 4;

QY 1 CEPLAELGTMDCTHHPGNFSFSSQCAFSCSEGTNLGIEETTCGPFGNSSPEPTCOV 60  
 DB 107 CTP---PPLISFVPISSALHFGSTVKYSCVGSEFLRGNTLTLCQDDGMSSPLPCVP 163  
 QY 61 IQC-EPLSAPDLGIMNCSDPLASFTSACTFTICSGTELIGKKTICSSGIMNSPPI 119  
 DB 164 VECQPEIIPN-GIID-----VOGLAYLSTALYTCKPGFELVGNNTTTLGNGHVLGKPT 218  
 QY 120 COKLD 124  
 DB 219 CKATE 223

RESULT 22  
 Q91YB6 PRELIMINARY; PRT; 1236 AA.  
 AC 091YB6;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE COMPLEMENT INHIBITORY FACTOR H.  
 GN FH.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=LIVER;  
 RA Demberg T., Goetze O., Schlat G.;  
 RT "Rat complement factor H: molecular cloning, sequencing and expression  
 RT in tissues and isolated cells";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ320522; CAC67513.1; -  
 SQ SEQUENCE 1236 AA; 140343 MW; 1AC89FPA28232EBF CRC64;

Query Match 23.1%; Score 169; DB 11; Length 1236;  
 Best Local Similarity 28.8%; Pred. No. 2.8e-11;  
 Matches 36; Conservative 26; Mismatches 47; Indels 16; Gaps 5;

QY 14 CTH----PFGN-----FSSQCAFSCSEGTNLG-IEETTCGPFGNSSPEPTCOV 60  
 DB 85 CGHGDPTFGSFLRVAEFGAKVYVTCDEGVLGEIDYRECDADG-WINDIPICEV 143  
 QY 61 IQCEPLSAPDLG--IMNCSDPLASFTSACTFTICSGTELIGKKTICSSGIMNSPPI 118  
 DB 144 VKCLPVELENGRIVSGAEPDDEYFGVYRFECNSGFKIESQKEMHCSGNGLSNKP 203  
 QY 119 ICOKL 123  
 DB 204 QCYEI 208

RESULT 23  
 Q96JUT7 PRELIMINARY; PRT; 1316 AA.  
 AC 096JUT7;  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE CDNA FLJ14964 FIS, CLONE PLACE4000581, MODERATELY SIMILAR TO  
 DE FIBROPELIN I PRECURSOR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-PLACENTA;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Aotsuka S., Yoshikawa Y.,



RA Matsunawa H., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A.,  
 RA Nakamura Y., Nagahari K., Masuko Y., Sasaki N.,  
 RT "NEBO human cDNA sequencing project";  
 RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
 DR EMBL AK021870; BAB55420.1; -  
 SQ SEQUENCE 1316 AA; 144524 MW; 80615BBA3A4F00A5 CRC64;

Query Match 23.0%; Score 168.5; DB 4; Length 1316;  
 Best Local Similarity 30.9%; Pred.No.3,4e-11;  
 Matches 42; Conservative 15; Mismatches 68; Indels 11; Gaps 4;

OY 1 CEPLAPELGTW---DCTHPGNGFSFSSQAFSCSEGTNLGTGTEETTCGPGWSSPEPT 57  
 DB 145 CALKLPPENGFYQNTC-----NNHFNACGYRCHGPGPLVSSITLCIPNGIMSGLESY 199  
 OY 58 CQVIOCEPLAPDLGIMNCSHPLASFTSACTFICSEGTETLIGKKRTICSSGIMSNPS 117  
 DB 200 CWRFCRPHLRQPKHGHISCS--TREMLYKTTCTLVACDEGYLEGSDKLTGCGNSQWDGPE 257  
 OY 118 PIC-QKLKSFMIKE 132  
 DB 258 PRCVERHCSTFQMPKD 273

RESULT 24

OY 029530 PRELIMINARY: PRT: 2014 AA.  
 AC 029530:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE COMPLEMENT RECEPTOR 1 (FRAGMENT).  
 GN CRL.  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94292799; PubMed=8021505;  
 RA Birmingham D.J., Shen X.P., Hourcade D., Nickells M.W., Atkinson J.P.;  
 RT "Primary sequence of an alternatively spliced form of CRL. Candidate  
 for the 75,000 M(r) complement receptor expressed on chimpanzee  
 erythrocytes";  
 RL J. Immunol. 153:691-700(1994).  
 DR EMBL L24920; AAS1438.1; -  
 DR HSSP; P08603; IHP1.  
 DR InterPro: IPR001424; SOD\_CU\_ZN.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR000834; Zn\_cathopept.  
 DR Pfam: PF00084; Sush1; 30.  
 DR SMART; SM00032; CCP; 30.  
 DR PROSITE; PS00133; CARBOXYPEPT\_ZN\_2; UNKNOWN\_2.  
 DR PROSITE; PS00087; SOD\_CU\_ZN\_1; UNKNOWN\_1.  
 KW Receptor.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 2014 AA; 221281 MW; 6D6C3A74D81FD1DB9 CRC64;

Query Match 22.9%; Score 167.5; DB 6; Length 2014;  
 Best Local Similarity 32.6%; Pred.No.7,4e-11;  
 Matches 46; Conservative 16; Mismatches 36; Indels 23; Gaps 6;

OY 1 CEPLAPELGTWCTHPGNGFSFSSQAFSCSEGTNLGTGTEETTCGPGFN--WSSPEPT 57  
 DB 1432 CGPPPEPFNGW--HINDTDFGSTVNVNSCNGEGRLLGSPSTTCLVSGNNVTWKKAPI 1488  
 OY 58 CQVIOCEP---ISAPDLGIMNCSHPLASFTSACTFICSEPT-----ELIGKKRTICE 108  
 DB 1489 CEIICEPEPTISNGDF---YSNNRASFRHNGTAVTYOCHTGPDEQTLFELVGERSIYCT 1544

OY 109 SS---GIWNSPITCOKIDK 125  
 DB 1545 SKDDQGVWSSPPRCISITNK 1565

RESULT 25

OY 016745 PRELIMINARY: PRT: 2039 AA.  
 AC 016745:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE COMPLEMENT RECEPTOR 1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94065175; PubMed=8245463;  
 RA Vik D.P., Wong W.W.;  
 RT "Structure of the gene for the F allele of complement receptor type 1  
 and sequence of the coding region unique to the S allele.";  
 RL J. Immunol. 151:6214-6224(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Vik D.P., Wong W.W.;  
 RL Submitted (JUN-1993) to the EMBL/Genbank/DBJ databases.  
 DR EMBL L17418; AAB60694.1; -  
 DR EMBL L17390; AAB60694.1; JOINED.  
 DR EMBL L17399; AAB60694.1; JOINED.  
 DR EMBL L17409; AAB60694.1; JOINED.  
 DR EMBL L17419; AAB60694.1; JOINED.  
 DR EMBL L17420; AAB60694.1; JOINED.  
 DR EMBL L17421; AAB60694.1; JOINED.  
 DR EMBL L17422; AAB60694.1; JOINED.  
 DR EMBL L17423; AAB60694.1; JOINED.  
 DR EMBL L17391; AAB60694.1; JOINED.  
 DR EMBL L17392; AAB60694.1; JOINED.  
 DR EMBL L17393; AAB60694.1; JOINED.  
 DR EMBL L17394; AAB60694.1; JOINED.  
 DR EMBL L17395; AAB60694.1; JOINED.  
 DR EMBL L17396; AAB60694.1; JOINED.  
 DR EMBL L17397; AAB60694.1; JOINED.  
 DR EMBL L17398; AAB60694.1; JOINED.  
 DR EMBL L17400; AAB60694.1; JOINED.  
 DR EMBL L17401; AAB60694.1; JOINED.  
 DR EMBL L17402; AAB60694.1; JOINED.  
 DR EMBL L17403; AAB60694.1; JOINED.  
 DR EMBL L17404; AAB60694.1; JOINED.  
 DR EMBL L17405; AAB60694.1; JOINED.  
 DR EMBL L17406; AAB60694.1; JOINED.  
 DR EMBL L17407; AAB60694.1; JOINED.  
 DR EMBL L17408; AAB60694.1; JOINED.  
 DR EMBL L17410; AAB60694.1; JOINED.  
 DR EMBL L17411; AAB60694.1; JOINED.  
 DR EMBL L17412; AAB60694.1; JOINED.  
 DR EMBL L17413; AAB60694.1; JOINED.  
 DR EMBL L17414; AAB60694.1; JOINED.  
 DR EMBL L17415; AAB60694.1; JOINED.  
 DR EMBL L17416; AAB60694.1; JOINED.  
 DR EMBL L17417; AAB60694.1; JOINED.  
 DR HSSP; P08603; IHP1.  
 DR InterPro: IPR001424; SOD\_CU\_ZN.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR000834; Zn\_cathopept.  
 DR Pfam: PF00084; Sush1; 30.  
 DR SMART; SM00032; CCP; 30.  
 DR PROSITE; PS00133; CARBOXYPEPT\_ZN\_2; UNKNOWN\_2.  
 DR PROSITE; PS00087; SOD\_CU\_ZN\_1; UNKNOWN\_1.  
 KW Receptor.  
 SQ SEQUENCE 2039 AA; 223603 MW; B82FCB11C6B16635 CRC64;



OY 101 GKKTICSSGIMNSPSPIC 120  
 DB 2779 GSSALTCMANGIMDRSLPKC 2798

## RESULT 28

O96RM4 PRELIMINARY; PRT; 3508 AA.  
 AC 096RM4:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CUB AND SUSHI MULTIPLE DOMAINS 1 PROTEIN.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21365705; PubMed=11472063;  
 RA Sun P.C., Uppaluri R., Schmidt A.P., Pashia M.E., Quant E.C.,  
 RA Sunwoo J.B., Gollin S.M., Scholnick S.B.;  
 RA "Transcript map of the 8p23 putative tumor suppressor region."; Genomics 75:17-25(2001).  
 RL EMBL: AF333704; AAK73475.1;  
 DR EMBL: AF333704; AAK73475.1;  
 SQ SEQUENCE 3508 AA; 382824 MW; 9268C3EBF3F78C18 CRC64;

Query Match 21.8%; Score 160; DB 4; Length 3508;  
 Best Local Similarity 26.4%; Pred. No. 1.1e-09;  
 Matches 37; Conservative 17; Mismatches 48; Indels 38; Gaps 4;

OY 5 EAPELGTMDCTHPFGN-----FSRSCAFSCSGTMTLTGIEETTCGPFQWSSP 54  
 DB 2673 QPVCVPTICGHP-GNPAHGTNGSEFLNDVNFCTNGYLGQVSRACRSNGQWSSP 2731  
 OY 55 EPTCOVIOCEPLSAPDLGIMNCSP-----LASFSFTSACTFICSEGTETI 100  
 DB 2732 LPTCRV-----NCSDPGFVENAIRHGOQNPESFEIGMSILYHCKKGFFLL 2778  
 OY 101 GKKTICSSGIMNSPSPIC 120  
 DB 2779 GSSALTCMANGIMDRSLPKC 2798

## RESULT 29

O921P0 PRELIMINARY; PRT; 390 AA.  
 AC 0921P0:  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE SIMILAR TO DECAV ACCELERATING FACTOR 1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC011314; AAH11314.1;  
 SQ SEQUENCE 390 AA; 42637 MW; 3FB68595F07A67FF CRC64;

Query Match 21.4%; Score 156.5; DB 11; Length 390;  
 Best Local Similarity 30.9%; Pred. No. 2.3e-10;  
 Matches 42; Conservative 16; Mismatches 63; Indels 15; Gaps 6;

OY 1 CEPLEAPELGTMDCTH---PFGNFSFSCAFSCSGTMTLTGIEETTCGPFQWSSP 54  
 DB 158 CRRKSCPFRKIDNGHINIPFG-ILFGSEINFSQNGYRLVGVSVFCSVTGNTVMDDE 216

OY 55 EPTCOVIOCEPLSAPDLGIMNCSP-----FSRSCAFSCSGTMTLTGIEETTCGPFQWSSP 54  
 DB 217 FVCTCHICPEPKINNGIMRGESD--SYTISQVYVYSCDKGVLYGNASTYCTVSKSV 274

OY 111 GIMNSPSPICQKLDKS 126  
 DB 275 GOWSSPPRC--IEKS 288

## RESULT 30

O08569 PRELIMINARY; PRT; 533 AA.  
 AC 008569:  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ACROSOMAL MATRIX COMPONENT AM67 PRECURSOR.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HARTLEY; TISSUE-TESTIS;  
 RX MEDLINE=97284752; PubMed=9139729;  
 RA Foster J.A., Friday B.B., Maulit M.T., Blobel C., Wintrey V.P.,  
 RA Olson G.E., Kim K.S., Gerton G.L.;  
 RA "AM67, a secretory component of the guinea pig sperm acrosomal matrix, is related to mouse sperm protein sp56 and the complement component 4-binding proteins."; J. Biol. Chem. 272:12714-12722(1997).  
 RL EMBL: U75654; AAC13888.1;  
 DR HSPF; P10998; 1YVD.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00084; sush1; 7.  
 DR SMART; SM00032; CCP; 7.  
 DR Signal.  
 FT SIGNAL 1 28 POTENTIAL.  
 FT CHAIN 29 533 ACROSOMAL MATRIX COMPONENT AM67.  
 SQ SEQUENCE 533 AA; 59772 MW; EDBDDE487A45389 CRC64;

Query Match 21.1%; Score 155; DB 11; Length 533;  
 Best Local Similarity 31.5%; Pred. No. 4.9e-10;  
 Matches 35; Conservative 20; Mismatches 42; Indels 14; Gaps 5;

OY 23 FSSQAFSCSGTMTLTGIEETTC--GPGNMSPPPTCOVIOCEPLSAPDLGIMNCSP 79  
 DB 110 FSGTIEFSCSKGSLIGSTTSCQSGKTVDMNDPLPECVIYKCD--SPPD--ISNGKHS 165  
 OY 80 LAS---FSTTACTFICSEGTETLGGKKTIC---ESSGIMNSPSPICQK 123  
 DB 166 GDEUDELTYGSLVTVCDPNYSILGNASISCLVANKTVGWSNPTCEKV 216

## RESULT 31

O93JA1 PRELIMINARY; PRT; 555 AA.  
 AC 093JA1:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE C4BP ALPHA-CHAIN PRECURSOR.  
 GN C4BPA.  
 OS Cavia porcellus (Guinea pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Hystriocognathi; Cavidae; Cavia.  
 OX NCBI\_TaxID=10141;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-EPIDIDYMIS, AND LIVER;  
 RX MEDLINE=21154058; PubMed=11254714;  
 RA Nonaka M.I., Wang G., Mori T., Okada H., Nonaka M.;



SEQUENCE 3564 AA; 387865 MW; 70824C55B0674609 CRC64;

Query Match 20.9%; Score 153; DB 11; Length 3564;

Best Local Similarity 26.7%; Pred. No. 7.7e-09;

Matches 35; Conservative 17; Mismatches 59; Indels 20; Gaps 4;

QY 1 CEPLAEPGLGMDCTHPFNG-----FSFSQCAFSCSGTNLTGIEETTCGPGFN 50

DB 2733 CVPI-----TCGHP-GNPAHGLTNGTEFNLNDLVNFTCTHGTGRLOQASRAQCRKSNQ 2783

QY 51 WSSPEPTCOVICPEPLSAPDLGIMNSHPL-ASFSTSACTFICSEGTGELIGKKTKICS 109

DB 2784 WSSPLTICRVNCSDDPSGVENAVRHQGNPPESFEVGTGVNWKCKGTYLLSGSALTCA 2843

QY 110 SGTWSNPSPIC 120

DB 2844 SGIMDRSLPKC 2854

RESULT 35

002839 PRELIMINARY; PRT; 363 AA.

AC 002839; PRELIMINARY; PRT; 363 AA.

DT 01-JUL-1997 (TREMBLrel. 04, Created)

DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PORCINE MEMBRANE COFACTOR PROTEIN.

OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.

NCBI\_TaxId=9823;

RN [1]

SEQUENCE FROM N.A.

RX MEDLINE=97343414; PubMed=9199970;

RA Toyomura K., Fujimura T., Murakami H., Natsume T., Shigenisa T.,

Inoue N., Takeda J., Kinoshita T.;

RT "Molecular cloning of a pig homologue of membrane cofactor protein

(CDP6).";

Int. Immunol. 9:869-876(1997).

DR EMBL; D70897; BAA20476.1; -.

DR HSSP; P10998; 1VD.

DR InterPro: IPR000436; Sush1\_SCR\_CCP.

DR Pfam: PF00084; sush1. 4.

DR SMART: SM00032; CCP; 4.

SEQUENCE 363 AA; 39692 MW; 8A5FF329B8E03153 CRC64;

Query Match 20.8%; Score 152.5; DB 6; Length 363;

Best Local Similarity 29.5%; Pred. No. 6.2e-10;

Matches 38; Conservative 16; Mismatches 64; Indels 11; Gaps 4;

QY 1 CEPLAEPGLGMDCTHPFNGFNSQCAFSCSGTNLTGIEETTCGPGFN--WSSPEPT 57

DB 108 CSNLPLPLNGQV--SYPNCDMLFGSKAQFTCNMGFIICAEYTCQVSGVMWMSPEPL 165

QY 58 COVIOCEPLSAPDLGIMNSHPLASFSTSACTFICSEGTGELIGKKTKICS 112

DB 166 CEKILCKPGEIIPNGKYTNHSHK-DVEYNEVYYSCLSTGDEFSLVGESSLFCIGKDE 224

QY 113 WSNPSPIC 121

DB 225 WSDPECK 233

RESULT 36

09J2M6 PRELIMINARY; PRT; 395 AA.

AC 09J2M6; PRELIMINARY; PRT; 395 AA.

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE COMPLEMENT BINDING PROTEIN.

OS Macaca mulatta rhadinovirus 26-95.

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Gammaherpesvirinae; Rhadinovirus.

NCBI\_TaxId=119193;

RN [1]

SEQUENCE FROM N.A.

RC STRAIN=MACACA MULATTA RHADINOVIRUS ISOLATE 26-95;

RX MEDLINE=20173730; PubMed=10708456;

RA Alexander L., Denekamp L., Knapp A., Auerbach M.R., Damania B.,

Destrofers R.C.;

RT "The primary sequence of rhesus monkey rhadinovirus isolate 26-95;

RT sequence similarities to Kaposi's sarcoma-associated herpesvirus and

RL J. Virol. 74:3388-3398(2000).

DR EMBL; AF210726; AAF59982.1; -.

DR HSSP; P10998; 1VD.

DR InterPro: IPR001230; Preyn1n.

DR Pfam: PF00084; sush1\_SCR\_CCP.

DR SMART: SM00032; CCP; 4.

DR PROSITE: PS00294; PRENYLATION; UNKNOWN\_1.

SEQUENCE 395 AA; 43922 MW; B4C9C6F2E226AE06 CRC64;

Query Match 20.8%; Score 152.5; DB 12; Length 395;

Best Local Similarity 31.8%; Pred. No. 6.9e-10;

Matches 35; Conservative 18; Mismatches 46; Indels 11; Gaps 5;

QY 21 FSFSQCAFSCSGTNLTGIEETTC-----GPGMNSPEPTCOVIOCEPLSAPDLGINN 75

DB 109 FRKSNITKCNMGVLLGATVTRCLKYDSNLVDMPAAPTCEIEKCR--KQPD--IEN 164

QY 76 CS-HPLASF-SFTSACTFICSEGTGELIGKKTKICSSGIMSNPSPIC 123

DB 165 GRYVPQVEFYNYLETITFCNKDFSLIGNTTTCMTGTWSSVPCQOI 214

RESULT 37

090611 PRELIMINARY; PRT; 354 AA.

AC 090611; PRELIMINARY; PRT; 354 AA.

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE FURROWED (FRAGMENT).

FW.

OS Drosophila ananassae (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxId=7217;

RN [1]

SEQUENCE FROM N.A.

RA Marsh B.J., Chen Y., Stephan W.;

RT "Interaction of Natural Selection and Gene Flow in Drosophila

RT ananassae.";

RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.

DR EMBL; AF185289; AAF03885.1; -.

DR Flybase; FBgn0014574; Dana\fw.

DR InterPro: IPR002396; Selectin.

DR InterPro: IPR000436; Sush1\_SCR\_CCP.

DR Pfam: PF00084; sush1. 6.

DR PRINTS; PR00343; SELECTIN.

DR SMART; SM00032; CCP; 6.

FT NON\_TER 1

FT NON\_TER 354

SEQUENCE 354 AA; 38526 MW; AD02699F25DFCAB CRC64;

Query Match 20.7%; Score 151.5; DB 5; Length 354;

Best Local Similarity 30.0%; Pred. No. 8e-10;

Matches 39; Conservative 19; Mismatches 57; Indels 15; Gaps 4;

```

OY 5 EAPBLGTMDCTHP-----FG-NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 55
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 151 EAPVCELYTCCEMPVPVPGSYVYDYDNVMSKIKYSQCDPHIMHG:SDLECLDSEGMSIDA 210
OY 56 PFCOYIOCEPLSAPDLGIMNCNHPLASFS--FTSACTFICSGTELIGKKTKICSSSGIWM 113
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 211 PFCSEYIDCGPI---LPIPYGGHKYVNTSYVGSSEVSCIOSHNLGCVKRCQCLESGVM 266
OY 114 SNMSPICOKL 123
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 267 SDASPKCEEI 276

RESULT 38
OY 09H4R2 PRELIMINARY; PRT; 410 AA.
ID 09H4R2
AC 09H4R2
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE BA524D16A.1 (SUSHI-REPEAT-CONTAINING PROTEIN) (FRAGMENT).
GN SRFX.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wilson S.;
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL391688; CAC16060.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF002494; HXR; 1.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
FT NON_TER 1
SQ SEQUENCE 410 AA; 46781 MW; D32228E1A1AFAF370 CRC64;

Query Match 20.7%; Score 151.5; DB 4; Length 410;
Best Local Similarity 29.7%; Pred. No. 9.4e-10;
Matches 35; Conservative 13; Mismatches 63; Indels 7; Gaps 3;

OY 10 GTMCTHPEFG---NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 66
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 12 GEATCYSPKGNVSHSLGTRELSCDRGFRIGRRSVQCLPSRRWSG-TAYCROMCHAL 70
OY 67 SAPDLGIMNCNHPLASFSFTSACTFICSEGTTELIGKKTKICSSSGIWMSPICOKLD 124
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 71 PFITSGYTCTNGVL---LDSRCDYSCSSGYHLEGDRSRICMEDGRNMSGGPVQVDID 125

RESULT 39
OY 060687 PRELIMINARY; PRT; 465 AA.
ID 060687
AC 060687
DT 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE SUSHI-REPEAT PROTEIN.
GN SRPUL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kurosawa H., Inukai T., Inaba T., Goi K., Chang K.-S., Shinjo T.,
RA Rakestraw K.M., Neve C.W., Look T.A.;
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

```

```

DR EMBL; AF060567; AAC15765.1; -.
DR InterPro; IPR001128; Cyt_P450.
DR InterPro; IPR000410; HXR.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF02494; HXR; 1.
DR Pfam; PF00084; sush1; 3.
DR SMART; SM00032; CCP; 3.
DR PROSITE; PS00086; CYTOCHROME_P450; UNKNOWN_1.
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187F3EFPB8 CRC64;

Query Match 20.7%; Score 151.5; DB 4; Length 465;
Best Local Similarity 29.7%; Pred. No. 1.1e-09;
Matches 35; Conservative 13; Mismatches 63; Indels 7; Gaps 3;

OY 10 GTMCTHPEFG---NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCOYIOCEPL 66
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 67 GEATCYSPKGNVSHSLGTRELSCDRGFRIGRRSVQCLPSRRWSG-TAYCROMCHAL 125
OY 67 SAPDLGIMNCNHPLASFSFTSACTFICSEGTTELIGKKTKICSSSGIWMSPICOKLD 124
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 126 PFITSGYTCTNGVL---LDSRCDYSCSSGYHLEGDRSRICMEDGRNMSGGPVQVDID 180

RESULT 40
OY 022328 PRELIMINARY; PRT; 560 AA.
ID 022328
AC 022328
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE HYPOHETICAL 61.6 KDA PROTEIN.
GN T07H6.5.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditioidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC SRRAIN-BRISTOL N2;
RX MEDLINE-99069613; PubMed-9851916;
RA None;
RT "Genome sequencing of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Geisel C.;
RT "The sequence of C. elegans cosmid T07H6.";
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL EMBL; U53344; AA86225.1; -.
DR HSSP; P10998; 1VVD.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 7.
DR SMART; SM00032; CCP; 7.
KW Hypothetical protein.
SQ SEQUENCE 560 AA; 61619 MW; FB8923BAC1B320C9 CRC64;

Query Match 20.5%; Score 150.5; DB 5; Length 560;
Best Local Similarity 26.6%; Pred. No. 1.8e-09;
Matches 33; Conservative 24; Mismatches 62; Indels 5; Gaps 2;

OY 1 CEPLPELGTMDCTHPEFG-NFSSSQCAFSCSEGTNLTGIEETTCGFGWSSPEPTCO 59
    ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 126 CKRACPDGDIENGIRREDTFFEPHHVYSCNPGFLVIGSTSRQCSSNGEWTNPANCK 185

```



**This Page Blank (uspto)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:23:43 ; Search time 18.65 Seconds

(without alignments)  
274.047 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328

Perfect score: 733

Sequence: 1 CEPLAPELLCTMDCTHPFGN.....WSNPSPICQKDKFSMIKE 132

Scoring table: BIOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SWISSPROT\_40:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	99.2	372	1	LEM1_HUMAN
2	727	99.2	372	1	LEM1_PANTR
3	724	98.8	372	1	LEM1_PONPY
4	690	94.1	372	1	LEM1_MACMU
5	690	94.1	372	1	LEM1_PAPHA
6	537	73.0	370	1	LEM1_BOVIN
7	535	73.0	372	1	LEM1_RAT
8	518	70.7	372	1	LEM1_MOUSE
9	315	43.0	830	1	LEM3_HUMAN
10	313	42.7	768	1	LEM3_RAT
11	305.5	41.7	612	1	LEM2_MOUSE
12	300	40.9	611	1	LEM2_CANFA
13	300	40.9	768	1	LEM3_MOUSE
14	300	40.9	769	1	LEM3_SHEEP
15	299.5	40.9	549	1	LEM2_RAT
16	296	40.4	610	1	LEM2_HUMAN
17	290	38.6	646	1	LEM3_BOVIN
18	288	39.3	551	1	LEM2_RABIT
19	271.5	37.0	485	1	LEM2_BOVIN
20	245	33.4	484	1	LEM2_PIG
21	178	24.3	958	1	HIG_DROME
22	173	23.6	1231	1	CFAH_HUMAN
23	163.5	22.3	2039	1	CRI_HUMAN
24	160	21.8	1234	1	CFAH_MOUSE
25	155.5	21.2	390	1	DAFL_MOUSE
26	155.5	21.2	1019	1	LFC_TACTR
27	150.5	20.5	1025	1	CR2_MOUSE
28	149.5	20.4	597	1	CABP_MOUSE
29	147.5	20.1	685	1	CFAH_BOVIN
30	146	19.9	558	1	CABP_RAT
31	145	19.8	610	1	CABP_BOVIN
32	144.5	19.7	469	1	CABP_MOUSE
33	144	19.6	258	1	CABP_RAT

34	142	19.4	507	1	DAF_CAVPO	Q60401 cavia porce
35	141.5	19.3	345	1	APDH_BOVIN	P17690 bos taurus
36	141	19.2	1033	1	CR2_HUMAN	P20023 homo sapien
37	140.5	19.2	764	1	CFAB_HUMAN	P00751 homo sapien
38	139	19.0	263	1	VCP_VACCV	P10998 vaccinia vi
39	139	19.0	381	1	DAF_HUMAN	P08174 homo sapien
40	137	18.7	340	1	DAF_PONPY	P49457 pongo pygma
41	136.5	18.6	377	1	MCP_HUMAN	P15529 homo sapien
42	135.5	18.5	1019	1	LFC_CARRO	Q26422 carcinoscor
43	134	18.3	1627	1	PAPA_HUMAN	Q13219 homo sapien
44	133.5	18.2	407	1	DAF2_MOUSE	Q61476 mus musculu
45	132.5	18.1	752	1	CO2_HUMAN	P06681 homo sapien

## ALIGNMENTS

RESULT ID	1	LEM1_HUMAN	STANDARD:	PRT:	372 AA.
AC	P14151: P15023:				
DT	01-JAN-1990 (Rel. 13, Created)				
DT	01-FEB-1991 (Rel. 17, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte surface antigen Leu-8) (TQ1) (gp90-MEL)				
DE	(Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).				
GN	SELL OR LYAM1 OR LNH1.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89315837; PubMed=2664786;				
RA	Siegelman M.H., Weissman I.L.;				
RT	"Human homologue of mouse lymph node homing receptor: evolutionary				
RT	conservation at tandem cell interaction domains.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 86:5562-5566(1989).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Tonsil;				
RX	MEDLINE=89310350; PubMed=2473156;				
RA	Camerling D., James S.P., Stamenkovic I., Seed B.;				
RT	"Leu-8/TQ1 is the human equivalent of the Mel-14 lymph node homing				
RT	receptor.";				
RL	Nature 342:78-82(1989).				
RN	[4]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=89308881; PubMed=2663882;				
RA	Bowen B.R., Nguyen T., Lasky L.A.;				
RT	"Characterization of a human homologue of the murine peripheral lymph				
RT	node homing receptor.";				
RL	J. Cell Biol. 109:421-427(1989).				
RN	[5]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=90243637; PubMed=1692315;				
RA	Ord D.C., Ernst T.J., Zhou L.J., Rambaldi A., Spertini O.,				
RT	Griffin J., Tedder T.F.;				
RT	"Structure of the gene encoding the human leukocyte adhesion				
RT	molecule-1 (TQ1, Leu-8) of lymphocytes and neutrophils.";				
RL	J. Biol. Chem. 265:7760-7767(1990).				
RN	[6]				
RP	3D-STRUCTURE MODELING.				

```

RX MEDLINE=96074584; PubMed=7488174;
RA Rajorath J., Aruffo A.;
RT "A template for generation and comparison of three-dimensional
RT selectin models."
RL Biochem. Biophys. Res. Commun. 216:1018-1023(1995).
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENUES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -1- DATABASE: NAME-PROW; NOTE=CD guide CD62L entry;
CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd62l.htm"
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M25280; AAC63053.1; -
DR EMBL; X16150; CAA34275.1; -
DR EMBL; X17519; CAB43536.1; -
DR EMBL; X17519; CAB43537.1; ALT_SRO.
DR EMBL; X16070; CAA34203.1; ALT_SRO.
DR EMBL; M32414; AAB60700.1; -
DR EMBL; M32406; AAB60700.1; JOINED.
DR EMBL; M32407; AAB60700.1; JOINED.
DR EMBL; M32408; AAB60700.1; JOINED.
DR EMBL; M32409; AAB60700.1; JOINED.
DR EMBL; M32410; AAB60700.1; JOINED.
DR EMBL; M32411; AAB60700.1; JOINED.
DR EMBL; M32412; AAB60700.1; JOINED.
DR EMBL; M32413; AAB60700.1; JOINED.
DR PIR; A33912; A33912.
DR PIR; A34015; A34015.
DR PIR; S06798; S06798.
DR PIR; J10104; J10104.
DR PDB; 1KJB; 03-APR-96.
DR GlycoSuiteDB; P14151; -.
DR MIM; 153240; -.
DR InterPro; IPR000561; EGF-like.
DR InterPro; IPR002396; Selectin.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR InterPro; IPR001304; lectin_c.
DR Pfam; PF00008; EGF; 1.
DR Pfam; PF00059; lectin_c; 1.
DR Pfam; PF00084; sush1; 2.
DR PRINTS; PRO0343; SELECTIN.
DR SMART; SM00032; CCP; 2.
DR SMART; SM00034; CLECT; 1.
DR SMART; SM00181; EGF; 1.
DR PROSITE; PS00022; EGF_1; 1.
DR PROSITE; PS01186; EGF_2; 1.
DR PROSITE; PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE; PS50041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat; 3d-structure.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 29 372
FT DOMAIN 39 332
FT TRANSMEM 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155

```

```

FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 311 311 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 37 37 D -> Y (IN REF. 4).
FT CONFLICT 178 178 Y -> H (IN REF. 4).
FT CONFLICT 193 193 F -> L (IN REF. 1 AND 4).
FT CONFLICT 213 213 P -> S (IN REF. 3).
FT CONFLICT 214 214 L -> F (IN REF. 4).
FT CONFLICT 218 220 SFS -> NEN (IN REF. 2).
FT CONFLICT 242 242 G -> R (IN REF. 2).
SQ SEQUENCE 372 AA; 42187 MW; 6EA918BCA2D3643 CRC64;

Query Match 99.2%; Score 727; DB 1; Length 372;
Best Local Similarity 99.2%; Pred. No. 7,5e-64;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAEPLGTMDCTHPEFGNFSSQCAFSCEGTNLGTIEETGCGPFGWMSSEPTQCV 60
DB 197 CEPLAEPLGTMDCTHPEFGNFSSQCAFSCEGTNLGTIEETGCGPFGWMSSEPTQCV 256
QY 61 IQCEPLAPDLGIMNCCHPLASFSTACPFICSEGLIGKKTTCGSSGIMNSPDC 120
DB 257 IQCEPLAPDLGIMNCCHPLASFSTACPFICSEGLIGKKTTCGSSGIMNSPDC 316
QY 121 QKLDKFSMIKE 132
DB 317 QKLDKFSMIKE 328

RESULT 2
LEMI_PANTR ID LEMI_PANTR STANDARD; PRT; 372 AA.
AC 095237;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion
DE molecule-1) (LAW-1) (Leukocyte-endothelial cell adhesion molecule 1)
DE (IECAM1) (CD62L).
DS SELL.
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
OX NCBI_TaxID=9598;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,
RA Teurshita N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBD databases.
CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL
CC VENUES IN PERIPHERAL LYMPH NODES.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -

```

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U73728; AAB18248.1; -  
DR HSSP: P14151; 1KJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00059; lectin\_g\_1.  
DR Pfam: PF00084; sushi\_2.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP\_2.  
DR SMART: SM00034; CLECT\_1.  
DR SMART: SM00181; EGF\_1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS00186; EGF\_2; 1.  
DR PROSITE: PS00615; C-type-LECTIN\_1; 1.  
DR PROSITE: PS00615; C-type-LECTIN\_2; 1.  
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28  
FT PROPEP 29 38  
FT CHAIN 39 372  
FT DOMAIN 39 332  
FT TRANSMEM 333 355  
FT DOMAIN 356 372  
FT DOMAIN 55 155  
FT DOMAIN 156 192  
FT DOMAIN 196 255  
FT DOMAIN 258 317  
FT DISULFID 57 155  
FT DISULFID 128 147  
FT DISULFID 160 171  
FT DISULFID 165 180  
FT DISULFID 182 191  
FT DISULFID 197 241  
FT DISULFID 227 254  
FT DISULFID 259 303  
FT DISULFID 289 316  
FT CARBOHD 60 60  
FT CARBOHD 104 104  
FT CARBOHD 177 177  
FT CARBOHD 216 216  
FT CARBOHD 232 232  
FT CARBOHD 246 246  
FT CARBOHD 271 271  
SQ SEQUENCE 372 AA; 42188 MW; 6EA991802A2D3643 CRC64;

Query Match 99.2%; Score 727; DB 1; Length 372;  
Best Local Similarity 99.2%; Pred. No. 7.5e-64;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 CEPLAPELGTMDCTHPGNGFSFSCAFSCSEGTNLNGIEFTTGGPFGNMSPEPTCOV 60  
DB 197 CEPLAPELGTMDCTHPGNGFSFSCAFSCSEGTNLNGIEFTTGGPFGNMSPEPTCOV 256  
OY 61 IOCEPLASAPDLGIMNCSPHPLASFSFTSACTFICSEGTLEIGKKKTYICSSGWSNPSPIC 120  
DB 257 IOCEPLASAPDLGIMNCSPHPLASFSFTSACTFICSEGTLEIGKKKTYICSSGWSNPSPIC 316  
OY 121 OKLDSFSMIKE 132  
DB 317 OKLDSFSMIKE 328

RESULT 3  
LEMI\_PONPY

ID LEMI\_PONPY STANDARD; PRT; 372 AA.  
AC 095235;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion  
DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)  
DE (LECAM1) (CD62L).  
GN SEL.  
OS Pongo pygmaeus (Orangutan).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pongo.  
OX NCBI\_TaxID=9600;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Buduan J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L.,  
RA Tsunashita N.;  
RL Submitted (NOV-1996) to the EMBL/Genbank/DBJ databases.  
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE  
CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
CC VENULES IN PERIPHERAL LYMPH NODES.  
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U73729; AAB18247.1; -  
DR HSSP: P14151; 1KJB.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00059; lectin\_g\_1.  
DR Pfam: PF00084; sushi\_2.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP\_2.  
DR SMART: SM00034; CLECT\_1.  
DR SMART: SM00181; EGF\_1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS00186; EGF\_2; 1.  
DR PROSITE: PS00615; C-type-LECTIN\_1; 1.  
DR PROSITE: PS00615; C-type-LECTIN\_2; 1.  
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 28  
FT PROPEP 29 38  
FT CHAIN 39 372  
FT DOMAIN 39 332  
FT TRANSMEM 333 355  
FT DOMAIN 356 372  
FT DOMAIN 55 155  
FT DOMAIN 156 192  
FT DOMAIN 196 255  
FT DOMAIN 258 317  
FT DISULFID 57 155  
FT DISULFID 128 147  
FT DISULFID 160 171  
FT DISULFID 165 180  
FT DISULFID 182 191  
FT DISULFID 197 241  
FT DISULFID 227 254  
FT DISULFID 259 303  
FT DISULFID 289 316

```

FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 216 216 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42118 MW; 6517DD2213FF15E CRC64;

Query Match 98.8%; Score 724; DB 1; Length 372;
Best Local Similarity 98.5%; Pred. No. 1.5e-63;
Matches 130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGPFSSQCAFSCSGTNTLGTIEETTCGPGWSSPEPTCOV 60
DB 197 CEPLAPDLGMDCTHPGPFSSQCAFSCSGTNTLGTIEETTCGPGWSSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTTCSSGIMSNPSPIC 120
DB 257 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTTCSSGIMSNPSPIC 316
QY 121 QKDKSFSMIKE 132
DB 317 QKDKSFSMIKE 328

RESULT 4
LEMI_MACMU STANDARD; PRT; 372 AA.
ID LEMI_MACMU
AC Q95198;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE GN
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RA Budman J.I., Fu H., Johnson C.E., Thakur A.B., Berg E.L., Tsunashita N.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC -!- SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -!- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U73730; AAB18246.1; -.
DR HSSP: P14151; AKJB.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR000436; Sushi_SCR_CCP.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00008; Egf_1.

```

```

DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF00084; sushi; 2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 2.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN_1; 1.
DR PROSITE: PS50041; C-TYPE LECTIN_2; 1.
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sushi; Repeat.
FT SIGNAL 1 28 BY SIMILARITY.
FT PROPEP 29 38 BY SIMILARITY.
FT CHAIN 39 372 L-SELECTIN.
FT DOMAIN 39 332 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 333 355 POTENTIAL.
FT DOMAIN 356 372 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 55 155 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 156 192 EGF-LIKE.
FT DOMAIN 196 255 SUSHI 1.
FT DOMAIN 258 317 SUSHI 2.
FT DISULFID 57 155 BY SIMILARITY.
FT DISULFID 128 147 BY SIMILARITY.
FT DISULFID 160 171 BY SIMILARITY.
FT DISULFID 165 180 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 197 241 BY SIMILARITY.
FT DISULFID 227 254 BY SIMILARITY.
FT DISULFID 259 303 BY SIMILARITY.
FT DISULFID 289 316 BY SIMILARITY.
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 104 104 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 226 226 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 246 246 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 372 AA; 42109 MW; 08BD7DD5AC549D6D CRC64;

Query Match 94.1%; Score 690; DB 1; Length 372;
Best Local Similarity 92.4%; Pred. No. 3e-60;
Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGPFSSQCAFSCSGTNTLGTIEETTCGPGWSSPEPTCOV 60
DB 197 CEPLAPDLGMDCTHPGPFSSQCAFSCSGTNTLGTIEETTCGPGWSSPEPTCOV 256
QY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTTCSSGIMSNPSPIC 120
DB 257 IQCEPLAPDLGIMNCSPHPLASFSTACTPFCSEGTLLGKRTTCSSGIMSNPSPIC 316
QY 121 QKDKSFSMIKE 132
DB 317 QKDKSFSMIKE 328

RESULT 5
LEMI_PAPHA STANDARD; PRT; 372 AA.
ID LEMI_PAPHA
AC Q28768;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
DE GN
OS Papio hamadryas (Hamadryas baboon).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Papio.

```

OX NCBI\_TaxID=9557;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97128794; PubMed-8973334;  
 RA "PCR cloning of the cDNA encoding baboon L-selectin."  
 RT Gene 181:219-220(1996).  
 RL  
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE  
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
 CC VENULES IN PERIPHERAL LYMPH NODES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U52074; AAB40903.1; -  
 CC HSSP: P14151; 1KJB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sushi\_2.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 2.  
 DR SMART: SM00181; EGF\_1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-type-LECTIN\_1; 1.  
 DR PROSITE: PS50041; C-type-LECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sushi; Repeat.  
 FT SIGNAL 1 28  
 FT PROPEP 29 38  
 FT CHAIN 39 372  
 FT DOMAIN 39 332  
 FT TRANSMEM 333 355  
 FT DOMAIN 356 372  
 FT DOMAIN 55 155  
 FT DOMAIN 156 192  
 FT DOMAIN 196 255  
 FT DOMAIN 258 317  
 FT DOMAIN 317 355  
 FT DISULFID 57 155  
 FT DISULFID 128 147  
 FT DISULFID 160 171  
 FT DISULFID 165 180  
 FT DISULFID 182 191  
 FT DISULFID 197 241  
 FT DISULFID 227 254  
 FT DISULFID 259 303  
 FT DISULFID 289 316  
 FT CARBOHYD 60 104  
 FT CARBOHYD 104 177  
 FT CARBOHYD 177 226  
 FT CARBOHYD 226 232  
 FT CARBOHYD 232 246  
 FT CARBOHYD 246 271  
 FT CARBOHYD 271 291  
 SQ SEQUENCE 372 AA; 42091 MW; 64E7BD5AC549D69 CRC64;

Query Match

94.1%; Score 690; DB 1; Length 372;

Best Local Similarity 92.4%; Pred. No. 3e-60;  
 Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 CEPLAEPELCTMDCTHPGNEFSFSSQCAFSCEGTNLGTIEETTCGPFGMSSPEPTQCV 60  
 Db 197 CEPLAEPELCTMDCTHPGNEFSFSSQCAFSCEGTNLGTIEETTCGPFGMSSPEPTQCV 256  
 QY 61 IOCEPLAPDLGIMNCNSHPLASFSTFSACTFSEGTGELGKRTTCESGIMSNPSPIC 120  
 Db 257 IOCEPLAPDLGIMNCNSHPLASFSTFSACTFSEGTGELGKRTTCESGIMSNPSPIC 316  
 QY 121 QKDRKFSMIKE 132  
 Db 317 QKDRKFSMIKE 328  
 RESULT 6  
 ID LEM1\_BOVIN STANDARD; PRT; 370 AA.  
 AC P98131;  
 DT 01-FEB-1996 (rel. 33, Created)  
 DT 01-FEB-1996 (rel. 33, Last sequence update)  
 DT 01-FEB-1996 (rel. 33, Last annotation update)  
 DE L-selectin precursor (Lymph node homing receptor) (Leukocyte adhesion  
 DE molecule-1) (LAM-1) (Leukocyte-endothelial cell adhesion molecule 1)  
 DE (LECAM1) (CD62L).  
 OS SELL.  
 GN Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-92164727; PubMed-1371468;  
 RA Walcheck B., White M., Kurk S., Kishimoto T.K., Jurila M.A.;  
 RT "Characterization of the bovine peripheral lymph node homing  
 RT receptor: a lectin cell adhesion molecule (LECAM).";  
 RL Eur. J. Immunol. 22:469-476(1992).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-9405503; PubMed-7694420;  
 RA Bosworth B.T., Dowbenko D., Shuster D.E., Harp J.A.;  
 RT "Bovine L-selectin: a peripheral lymphocyte homing receptor.";  
 RL Vet. Immunol. Immunopathol. 37:201-215(1993).  
 CC -1- FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATES THE ADHERENCE  
 CC OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL  
 CC VENULES IN PERIPHERAL LYMPH NODES.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X62882; CAA44676.1; -  
 CC HSSP: P14151; 1KJB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF\_1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sushi\_2.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 2.



```
OY 1 CEPLPAPLGTMDCTHPGNEFSFSSQCAFSCSEGTNLGIEFTTCGPGNMSPEPTQY 60
DB 197 CEPLPAPLGTMDCTHPGNEFSFSSQCAFSCSEGTNLGIEFTTCGPGNMSPEPTQY 256
OY 61 IQCEPLAPDGLGIMNCMHPLASFSTFSACTFTCSGTELLIGKKTICSSGIGWNSPSPIC 120
DB 257 IQCEPLAPDGLGIMNCMHPLASFSTFSACTFTCSGTELLIGKKTICSSGIGWNSPSPIC 316
OY 121 OKLDSFSMIKE 132
DB 317 OKTRRSFSKIKE 328

RESULT 8
LEML_MOUSE STANDARD: PRT: 372 AA.
AC P16337.
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE L-selectin precursor (lymph node homing receptor) (Leukocyte adhesion molecule-1) (LAM-1) (LY-22) (lymphocyte surface MEL-14 antigen) (Leukocyte-endothelial cell adhesion molecule 1) (LECAM1) (CD62L).
GN SELT OR LNHOR OR LY-22.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA SIEGELMAN M.H., van de RIJN M., Weissman I.L.;
RT "Mouse lymph node homing receptor cDNA clone encodes a glycoprotein revealing tandem interaction domains."
RL Science 243:1165-1172(1989).
RN [2]
RP SEQUENCE FROM N.A.
RA SIEGELMAN M.H., Cheng I.C., Weissman I.L., Wakeland E.K.;
RT "The mouse lymph node homing receptor is identical with the endothelial binding."
RL Cell 61:611-622(1990).
RN [3]
RP SEQUENCE FROM N.A.
RA MEDLINE=89168433; PubMed=2647302;
RA Laaky L.A., Singer M.S., Yednock T.A., Dowbenko D., Fennie C., Rodriguez H., Nguyen T., Stachel S., Rosen S.D.;
RT "Cloning of a lymphocyte homing receptor reveals a lectin domain."
RL Cell 56:1045-1055(1989).
RN [4]
RP SEQUENCE OF 1-360 FROM N.A.
RA MEDLINE=91169529; PubMed=2004776;
RA Dowbenko D.J., Diep A., Taylor B.A., Lusis A.J., Lasky L.A.;
RT "Characterization of the murine homing receptor gene reveals correspondence between protein domains and coding exons."
RL Genomics 9:270-277(1991).
CC -1 FUNCTION: CELL SURFACE ADHESION PROTEIN. MEDIATE THE ADHERENCE OF LYMPHOCYTES TO ENDOTHELIAL CELLS OF HIGH ENDOTHELIAL VENULES IN PERIPHERAL LYMPH NODES.
CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMS.
CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC -1 SIMILARITY: CONTAINS 2 SUSHI (SCR) DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
```

```
CC -----
DR EMBL: X14772; CAA32880.1; -
DR EMBL: M36005; AAA39722.1; -
DR EMBL: M36058; AAA39723.1; -
DR EMBL: M25324; AAA39431.1; -
DR EMBL: M64549; AAA75651.1; -
DR EMBL: M64440; AAA75651.1; JOINED.
DR EMBL: M64545; AAA75651.1; JOINED.
DR EMBL: M64548; AAA75651.1; JOINED.
DR PIR: A32375; A32375.
DR HSSP: P1451; 1KCB.
DR MGI: MGI:98279; Sel1.
DR InterPro: IPR000561; EGF-like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR001212; Somatomedin_B.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR InterPro: IPR001304; lectin_C.
DR Pfam: PF00008; EGF_1.
DR Pfam: PF00059; lectin_C_1.
DR Pfam: PF00084; sush1_2.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP_2.
DR SMART: SM00034; CLECT_1.
DR SMART: SM00181; EGF_1.
DR SMART: SM00201; SO_1.
DR PROSITE: PS00022; EGF_1_1.
DR PROSITE: PS01186; EGF_2_1.
DR PROSITE: PS00615; C-TYPE-LECTIN_1_1.
DR PROSITE: PS50041; C-TYPE-LECTIN_2_1.
DR Cell adhesion: Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW Selectin; Signal; Sush1; Repeat.
FT SIGNAL 1 28
FT PROPEP 29 38
FT CHAIN 39 372
FT DOMAIN 39 332
FT TRANSMEM 333 355
FT DOMAIN 356 372
FT DOMAIN 55 155
FT DOMAIN 156 192
FT DOMAIN 196 255
FT DOMAIN 258 317
FT DISULFID 57 155
FT DISULFID 128 147
FT DISULFID 160 171
FT DISULFID 165 180
FT DISULFID 182 191
FT DISULFID 197 241
FT DISULFID 227 254
FT DISULFID 229 303
FT DISULFID 289 316
FT CARBOHYD 60 60
FT CARBOHYD 104 104
FT CARBOHYD 177 177
FT CARBOHYD 216 216
FT CARBOHYD 226 226
FT CARBOHYD 246 246
FT CARBOHYD 278 278
FT CARBOHYD 288 288
FT CARBOHYD 308 308
FT CARBOHYD 320 320
FT CONFLICT 32 32
FT SEQUENCE 372 AA; 42288 MW; 4433EDF6ACB2B78 CMC64;
-----
Query Match 70.7%; Score 518; DB 1; Length 372;
Best Local Similarity 68.2%; Pred. No. 1.8e-43;
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;
OY 1 CEPLPAPLGTMDCTHPGNEFSFSSQCAFSCSEGTNLGIEFTTCGPGNMSPEPTQY 60
DB 197 CEPLPAPLGTMDCTHPGNEFSFSSQCAFSCSEGTNLGIEFTTCGPGNMSPEPTQY 256
OY 61 IQCEPLAPDGLGIMNCMHPLASFSTFSACTFTCSGTELLIGKKTICSSGIGWNSPSPIC 120
```

```

Db      257 VQCEPLAPELLGTMDCHPLGNFSFQSKCAPNCSEGRRELLGTARTQCGASGNWMSPPPIC 316
QY      121 QKDKSFSMAKE 132
       1 : : : : :
Db      317 QETNRSFSKIKE 328

RESULT 9
LEK3_HUMAN
ID      LEM3_HUMAN      STANDARD;      PRT;      830 AA.
AC      P16109;
DT      01-APR-1990 (Rel. 14, Created)
DT      01-AUG-1990 (Rel. 15, Last sequence update)
DT      16-OCT-2001 (Rel. 40, Last annotation update)
DE      P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)
DE      (CD62P) (Leukocyte-endothelial cell adhesion molecule 3) (LECAM3).
GN      SLIP OR GMRP.
OS      Homo sapiens (Human).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX      NCBI_TaxID=9606;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89168432; Pubmed=2466574;
RA      Johnston G.I., Cook R.G., McEver R.P.;
RT      Cloning of GMP-140, a granule membrane protein of platelets and
RT      endothelium: sequence similarity to proteins involved in cell
RT      adhesion and inflammation.";
RL      Cell 56:1033-1044(1989).
RN      [2]
RP      PALMITOYLATION.
RX      MEDLINE=93266599; Pubmed=7684381;
RA      Fujimoto T., Stroud E., Whately R.E., Prescott S.M., Muszbek L.,
RA      Laposta M., McEver R.P.;
RT      P-selectin is acylated with palmitic acid and stearic acid at
RT      cysteine 766 through a thioester linkage.";
RL      J. Biol. Chem. 268:11394-11400(1993).
RN      [3]
RP      STRUCTURE BY NMR OF 160-199.
RX      MEDLINE=97057176; Pubmed=8901515;
RA      Freedman S.J., Sanford D.G., Bachovchin W.W., Furie B.C., Baleja J.D.,
RA      Furie B.;
RT      Structure and function of the epidermal growth factor domain of P-
RT      selectin.";
RL      Biochemistry 35:13733-13744(1996).
RN      [4]
RP      3D-STRUCTURE MODELING OF 42-161.
RX      MEDLINE=94093388; Pubmed=7505680;
RA      Bajorath J., Stenkamp R., Arnuffo A.;
RT      "Knowledge-based model building of proteins: concepts and examples.";
RL      Protein Sci. 2:1798-1810(1993).
RN      [5]
RP      VARIANTS ASN-331; ASP-603; VAL-640 AND PRO-756.
RX      MEDLINE=98334547; Pubmed=9668170;
RA      Hermann S.M., Ricard S., Nicaud V., Mallet C., Evans A.,
RA      Rudaevets J.B., Arweiler D., Luc G., Cambien F.;
RT      The P-selectin gene is highly polymorphic: reduced frequency of the
RT      pro715 allele carriers in patients with myocardial infarction.";
RL      Hum. Mol. Genet. 7:1277-1284(1998).
CC      -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS
CC      TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
CC      INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
CC      LEUKOCYTES. THE LIGAND RECOGNIZED IS STAYL-Lewis X.
CC      -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC      -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS
CC      AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL
CC      ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO
CC      THE CELL SURFACE.
CC      -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.
CC      -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC      -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC      -1- SIMILARITY: CONTAINS 9 SUSHI (SCR) DOMAINS.

```

```

CC      -1- DATABASE: NAME=PROW; NOTE=CD guide CD62P entry;
CC      WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62p.htm".
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation
CC      of the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC      or send an email to license@isb-slb.ch).
CC      -----
DR      EMBL; M60234; AAA35910.1; -
DR      EMBL; M60217; AAA35910.1; JOINED.
DR      EMBL; M60218; AAA35910.1; JOINED.
DR      EMBL; M60219; AAA35910.1; JOINED.
DR      EMBL; M60222; AAA35910.1; JOINED.
DR      EMBL; M60223; AAA35910.1; JOINED.
DR      EMBL; M60224; AAA35910.1; JOINED.
DR      EMBL; M60225; AAA35910.1; JOINED.
DR      EMBL; M60226; AAA35910.1; JOINED.
DR      EMBL; M60227; AAA35910.1; JOINED.
DR      EMBL; M60228; AAA35910.1; JOINED.
DR      EMBL; M60229; AAA35910.1; JOINED.
DR      EMBL; M60231; AAA35910.1; JOINED.
DR      EMBL; M60232; AAA35910.1; JOINED.
DR      EMBL; M60233; AAA35910.1; JOINED.
DR      EMBL; M25322; AAA35911.1; -
DR      PIR; A30359; A30359.
DR      PDB; 1FSB; 01-APR-97.
DR      PDB; 1KUD; 03-APR-96.
DR      MIM; 173610; -
DR      InterPro; IPR000561; EGF-like.
DR      InterPro; IPR002396; Selectin.
DR      InterPro; IPR000436; Sushi_SCR_CCP.
DR      InterPro; IPR001304; Lectin_C.
DR      Pfam; PF00008; EGF_1.
DR      Pfam; PF00059; Lectin_C_1.
DR      Pfam; PF00084; sushi_9.
DR      PRINTS; PR00343; SELECTIN.
DR      SMART; SM00032; CCP; 9.
DR      SMART; SM00034; CLECT; 1.
DR      SMART; SM00181; EGF_1.
DR      PROSITE; PS00022; EGF_1; 1.
DR      PROSITE; PS01186; EGF_2; 1.
DR      PROSITE; PS00615; C-TYPE_LLECTIN_1; 1.
DR      PROSITE; PS50041; C-TYPE_LLECTIN_2; 1.
KW      Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;
KW      Selectin; Signal; Sushi; Repeat; Lipoprotein; Palmitate; Polymorphism;
KW      3D-structure.
FT      CHAIN 1 41
FT      DOMAIN 42 830
FT      TRANSMEM 772 795
FT      DOMAIN 796 830
FT      DOMAIN 158 158
FT      DOMAIN 159 195
FT      DOMAIN 199 258
FT      DOMAIN 261 320
FT      DOMAIN 323 382
FT      DOMAIN 385 444
FT      DOMAIN 447 506
FT      DOMAIN 509 568
FT      DOMAIN 571 630
FT      DOMAIN 641 700
FT      DOMAIN 703 762
FT      DISULFID 60 158
FT      DISULFID 131 150
FT      DISULFID 163 174
FT      DISULFID 168 183
FT      DISULFID 185 194
FT      DISULFID 200 244
FT      DISULFID 230 257
FT      DISULFID 262 306

```



```
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
FT DISULFID 416 443 BY SIMILARITY.
FT DISULFID 448 492 BY SIMILARITY.
FT DISULFID 478 505 BY SIMILARITY.
FT DISULFID 510 554 BY SIMILARITY.
FT DISULFID 540 567 BY SIMILARITY.
FT DISULFID 572 616 BY SIMILARITY.
FT DISULFID 602 629 BY SIMILARITY.
FT DISULFID 642 686 BY SIMILARITY.
FT DISULFID 672 699 BY SIMILARITY.
FT DISULFID 704 748 BY SIMILARITY.
FT DISULFID 734 761 BY SIMILARITY.
FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 411 411 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 460 460 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 518 518 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 716 716 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 723 723 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 741 741 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT LIPID 807 807 PALMITATE.
FT SITE 818 821 ENDOCYTOSIS SIGNAL (PROBABLE).
FT VARIANT 331 331 S -> N.
FT VARIANT 603 603 /FTID-VAR_004192.
FT VARIANT 640 640 N -> D.
FT VARIANT 640 640 /FTID-VAR_004193.
FT VARIANT 756 756 L -> V.
FT VARIANT 756 756 /FTID-VAR_004194.
T -> P (REDUCED FREQUENCY IN PATIENTS
WITH MYOCARDIAL INFARCTION).
/FTID-VAR_004195.
SEQUENCE 830 AA: 90844 MW: FBC407BA2579F6EB CMC64;
```

```
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG:
RX MEDLINE-94333817; PubMed-7520013;
RA Auchampach J.A., Oliver M.G., Anderson D.C., Manning A.M.;
RT "Cloning, sequence comparison and in vivo expression of the gene
RL encoding rat P-selectin.";
Gene 145:251-255(1994).
CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS
TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE
INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH
LEUKOCYTES. THE LIGAND RECOGNIZED IS STALTY-LEWIS X.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXAMINED: SPLEEN,
LUNG, BRAIN, LIVER, HEART, KIDNEY, THYMUS, SMALL INTESTINE.
CC -1- INDUCTION: BY ACUTE INFLAMMATION (PROBABLE).
CC -1- SIMILARITY: TO OTHER SELECTINS/LECTINS.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS; RAT P-LECTIN LACKS THE
HUMAN SUSHI-2 EQUIVALENT.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@sib-sib.ch).
-----
DR EMBL: L23088; AAA60325.1; -.
DR HSSP: P16109; IFSB.
DR InterPro: IPR000561; EGF_1like.
DR InterPro: IPR002396; Selectin.
DR InterPro: IPR00436; Sush1_SCR_CCP.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF00084; sush1; 8.
DR PRINTS: PR00343; SELECTIN.
DR SMART: SM00032; CCP; 8.
DR SMART: SM00034; CLECT; 1.
DR SMART: SM00181; EGF; 1.
DR PROSITE: PS00022; EGF_1; 1.
DR PROSITE: PS01186; EGF_2; 1.
DR PROSITE: PS00615; C-TYPE LECTIN; 1.
DR PROSITE: PS50041; C-TYPE LECTIN; 2; 1.
DR Cell adhesion; transmembrane; Glycoprotein; EGF-like domain; Lectin;
Selectin; Signal; Sush1; Repeat; Lipoprotein; Palmitate.
KW CHAIN 1
FT SIGNAL 41
FT TRANSMEM 42 768
FT DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 710 733 POTENTIAL.
FT DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).
FT DOMAIN 159 195 EGF-LIKE.
FT DOMAIN 199 258 SUSHI 1.
FT DOMAIN 261 320 SUSHI 2.
FT DOMAIN 323 382 SUSHI 3.
FT DOMAIN 385 444 SUSHI 4.
FT DOMAIN 447 506 SUSHI 5.
FT DOMAIN 509 568 SUSHI 6.
FT DOMAIN 579 638 SUSHI 7.
FT DOMAIN 641 700 SUSHI 8.
FT DISULFID 60 158 BY SIMILARITY.
FT DISULFID 131 150 BY SIMILARITY.
FT DISULFID 168 183 BY SIMILARITY.
FT DISULFID 185 194 BY SIMILARITY.
FT DISULFID 200 244 BY SIMILARITY.
FT DISULFID 230 257 BY SIMILARITY.
FT DISULFID 262 306 BY SIMILARITY.
FT DISULFID 292 319 BY SIMILARITY.
FT DISULFID 324 368 BY SIMILARITY.
FT DISULFID 354 381 BY SIMILARITY.
FT DISULFID 386 430 BY SIMILARITY.
```

FT	DISULFID	416	443	BY SIMILARITY.
FT <th>DISULFID</th> <th>448</th> <th>492</th> <th>BY SIMILARITY.</th>	DISULFID	448	492	BY SIMILARITY.
FT <th>DISULFID</th> <th>478</th> <th>503</th> <th>BY SIMILARITY.</th>	DISULFID	478	503	BY SIMILARITY.
FT <th>DISULFID</th> <th>510</th> <th>554</th> <th>BY SIMILARITY.</th>	DISULFID	510	554	BY SIMILARITY.
FT <th>DISULFID</th> <th>540</th> <th>567</th> <th>BY SIMILARITY.</th>	DISULFID	540	567	BY SIMILARITY.
FT <th>DISULFID</th> <th>580</th> <th>624</th> <th>BY SIMILARITY.</th>	DISULFID	580	624	BY SIMILARITY.
FT <th>DISULFID</th> <th>610</th> <th>637</th> <th>BY SIMILARITY.</th>	DISULFID	610	637	BY SIMILARITY.
FT <th>DISULFID</th> <th>642</th> <th>686</th> <th>BY SIMILARITY.</th>	DISULFID	642	686	BY SIMILARITY.
FT <th>DISULFID</th> <th>672</th> <th>699</th> <th>BY SIMILARITY.</th>	DISULFID	672	699	BY SIMILARITY.
FT <th>CARBOHYD</th> <th>45</th> <th>45</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	45	45	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>54</th> <th>54</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	54	54	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>107</th> <th>107</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	107	107	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>212</th> <th>212</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	212	212	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>347</th> <th>347</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	347	347	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>456</th> <th>456</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	456	456	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>603</th> <th>603</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	603	603	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>654</th> <th>654</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	654	654	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>661</th> <th>661</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	661	661	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>CARBOHYD</th> <th>679</th> <th>679</th> <th>N-LINKED (GLCNAC. . .) (POTENTIAL.)</th>	CARBOHYD	679	679	N-LINKED (GLCNAC. . .) (POTENTIAL.)
FT <th>LIPID</th> <th>745</th> <th>745</th> <th>PALMITATE (BY SIMILARITY).</th>	LIPID	745	745	PALMITATE (BY SIMILARITY).
FT <th>STATE</th> <th>756</th> <th>759</th> <th>ENDOTOXIN SIGNAL (PROBABLE).</th>	STATE	756	759	ENDOTOXIN SIGNAL (PROBABLE).
90 <th>SEQUENCE</th> <th>768 AA:</th> <th>83517 MM:</th> <th>26PD7EAS53FL316 CRC64:</th>	SEQUENCE	768 AA:	83517 MM:	26PD7EAS53FL316 CRC64:

Query Match	42.7%	Score 313;	DB 1;	Length 768;
Best Local Similarity	43.9%	Pred. No. 3.7e-23;		
Matches 54; Conservative	17;	Mismatches 52;	Indels 0;	Gaps 0;

```

QY      1  CEPLEAPLPGMDCTHPHPCGNSFFSQCACFSSCSECTNLTLGTEETTCGFGNMSSELPYCOV 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      262  CQSLSEAPLHGTMDCTHPLAAFAAYVSSCKFEQCPYRNRGSDILHCTDSGQMSSELPICEA 321
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      61  IQCEPLASAPDLGIMNCSHPLASFSFTSACPTTCESEGTLELLGKKTITESSGIMNSPSPIC 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      322  IACEPLSEPLHGSMDCFPSTCAFGYNSSCTFRCTEGVILMGNDAIHCADILGQMTAPAVC 381
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      121  QKL 123
      : : : : :
Db      382  EAL 384
      : : : : :

```

RESULT	11
LENM2_MOUSE	
ID	LENM2_MOUSE
AC	000690;
DT	01-APR-1993 (Rel. 25, Created)
DT	01-APR-1993 (Rel. 25, Last sequence update)
DT	01-FEB-1996 (Rel. 33, Last annotation update)
DE	E-selectin precursor (endothelial leukocyte adhesion molecule 1)
DE	(ELAM-1) (leukocyte-endothelial cell adhesion molecule 2) (LECAM2)
DE	(CD62E)
GN	SELE OR ELAM-1.
OS	Mus musculus (Mouse).
OC	Eukaryota; Eutheraa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murineae; Mus.
OX	NCBI_TaxID=10090;
RN	[1]
RX	SEQUENCE FROM N.A.
RX	MEDLINE=92283265; PubMed=1375914;
RA	Becker-Andre M., van haltsdultjnen R.H., Losberger C., wheilan T.,
RA	DeLaunay J.F.;
RT	"Murine endothelial leukocyte-adhesion molecule 1 is a close
RT	structural and functional homologue of the human protein.";
RL	Eur. J. Biochem. 206:401-411(1992).
RL	[2]
RP	SEQUENCE FROM N.A.
RP	MEDLINE=92340571; PubMed=1376846;
RA	Weller A., Isenmann S., Vestweber D.;
RA	"Cloning of the mouse endothelial selectins. Expression of both E
RT	and P-selectin is inducible by tumor necrosis factor alpha.";
RL	J. Biol. Chem. 267:15176-15183(1992).
CC	-1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND
CC	MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY



DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)  
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).  
 GN NEP OR GRMP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92340571; PubMed=1378846;  
 RA Weiler A., Iseemann S., Vestweber D.;  
 RT Cloning of the mouse endothelial selectins. Expression of both E-  
 RT and P-selectin is inducible by tumor necrosis factor alpha.";  
 RT J. Biol. Chem. 267:15176-15183(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92345617; PubMed=1379089;  
 RA Sanders W.E. Jr., Wilson R.W., Ballantyne C.M., Beaudet A.L.;  
 RT Molecular cloning and analysis of in vivo expression of murine P-  
 RT selectin.";  
 RT Blood 80:795-800(1992).  
 CC -1- FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS  
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE  
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH  
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.  
 CC -1- SUBCELLULAR LOCATION: TYPE I membrane protein.  
 CC -1- TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS  
 CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL  
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO  
 CC THE CELL SURFACE.  
 CC -1- INDUCTION: BY TNF-ALPHA.  
 CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1- SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS; MOUSE P-LECTIN LACKS  
 CC THE HUMAN SUSHI-2 EQUIVALENT.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M7861; AAA40008.1; -;  
 DR EMBL: M72332; AAA37712.1; -;  
 DR PIR: A42755; A42755.  
 DR HSSP: P16109; IFSB.  
 DR MGD: MGI:98280; Selp.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR000742; EGF 2.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1-SCR\_CCP.  
 DR InterPro: IPR001304; lectin\_c.  
 DR Pfam: PF00008; EGF; 1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR Pfam: PF00084; sush1; 8.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 8.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE-LECTIN\_1; 1.  
 DR PROSITE: PS00615; C-TYPE-LECTIN\_2; 1.  
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sush1; Repeat; Lipoprotein; Palmitate.  
 FT SIGNAL 1 41  
 FT CHAIN 42 768 P-SELECTIN.  
 FT DOMAIN 42 709 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 710 733 POTENTIAL.

FT DOMAIN 734 768 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 58 158 C-TYPE LECTIN (SHORT FORM).  
 FT DOMAIN 159 195 EGF-LIKE.  
 FT DOMAIN 199 258 SUSHI 1.  
 FT DOMAIN 261 320 SUSHI 2.  
 FT DOMAIN 323 382 SUSHI 3.  
 FT DOMAIN 385 444 SUSHI 4.  
 FT DOMAIN 447 506 SUSHI 5.  
 FT DOMAIN 509 568 SUSHI 6.  
 FT DOMAIN 579 638 SUSHI 7.  
 FT DOMAIN 641 700 SUSHI 8.  
 FT DISULFID 60 158 BY SIMILARITY.  
 FT DISULFID 131 150 BY SIMILARITY.  
 FT DISULFID 163 174 BY SIMILARITY.  
 FT DISULFID 168 183 BY SIMILARITY.  
 FT DISULFID 185 194 BY SIMILARITY.  
 FT DISULFID 200 244 BY SIMILARITY.  
 FT DISULFID 230 257 BY SIMILARITY.  
 FT DISULFID 262 306 BY SIMILARITY.  
 FT DISULFID 292 319 BY SIMILARITY.  
 FT DISULFID 324 368 BY SIMILARITY.  
 FT DISULFID 354 381 BY SIMILARITY.  
 FT DISULFID 386 430 BY SIMILARITY.  
 FT DISULFID 416 443 BY SIMILARITY.  
 FT DISULFID 448 492 BY SIMILARITY.  
 FT DISULFID 478 505 BY SIMILARITY.  
 FT DISULFID 510 554 BY SIMILARITY.  
 FT DISULFID 540 567 BY SIMILARITY.  
 FT DISULFID 580 624 BY SIMILARITY.  
 FT DISULFID 610 637 BY SIMILARITY.  
 FT DISULFID 642 686 BY SIMILARITY.  
 FT DISULFID 672 699 BY SIMILARITY.  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 603 603 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 654 654 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 661 661 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 679 679 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT LIPID 745 745 PALMITATE (BY SIMILARITY).  
 FT SITE 756 759 ENDOCYTOSIS SIGNAL (PROBABLE).  
 FT CONFLICT 724 724 A -> E (IN REF. 2).  
 FT SEQUENCE 768 AA; 83098 MW; E5173074D2F6568 CRC64;  
 SQ  
 Query Match 40.9%; Score 300; DB 1; Length 768;  
 Best Local Similarity 42.3%; Pred. No. 6; 9e-22;  
 Matches 52; Conservative 17; Mismatches 54; Indels 0; Gaps 0;  
 QY 1 CEPLAEPLGTMCTHFGNPSFSCAFSCSESTNTLTGIEFTTCGPFQWSSPEPTQCV 60  
 DB 200 CGKVNIPQHVLMNCSHPLGEFSFSQCTFSCABEYELDGELOCLASGIWNNPPKCD 259  
 QY 61 IOCEPLSAPDLGIMNCSPHPLASFSTFACFTFICSEGTETIGKKKTCSSGIMNSPIC 120  
 DB 260 VQCCSLAPHPHGMACHMPILAFAYDSSCKFEQCPGRANGSNMTHOTSGGQMSPEPLPTC 319  
 QY 121 QKL 123  
 DB 320 EAI 322  
 RESULT 14  
 LEM3\_SHEEP STANDARD; PRT; 769 AA.  
 ID LEM3\_SHEEP  
 AC P98109;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE P-selectin precursor (Granule membrane protein 140) (GMP-140) (PADGEM)  
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).  
 GN SELP.  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

CC Bovidae; Caprinae; Ovis.  
 OX NCBI\_TaxID=99940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RA Burns S.A., Neufeld E.J., Donady J.J.;  
 RL Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MYELOID CELLS THAT BINDS  
 TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE  
 INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH  
 LEUKOCYTES. THE LIGAND RECOGNIZED IS SIALYL-LEWIS X.  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1 SIMILARITY: CONTAINS 1 EGF-TYPE LECTIN FAMILY DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 8 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; L34270; AAB59261.1; -.  
 DR HSSP; P16109; IFSB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR001304; Lectin\_c.  
 DR Pfam; PF00008; EGF\_1.  
 DR Pfam; PF00059; Lectin\_c; 1.  
 DR Pfam; PF00084; Sush1; 8.  
 DR PRINTS; PR00343; SELECTIN.  
 DR SMART; SM00034; CLECT; 1.  
 DR SMART; SM00181; EGF; 1.  
 DR PROSITE; PS00022; EGF\_1; 1.  
 DR PROSITE; PS01186; EGF\_2; 1.  
 DR PROSITE; PS00615; C-type-LECTIN\_1; 1.  
 DR PROSITE; PS00615; C-type-LECTIN\_2; 1.  
 DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sush1; Repeat.  
 FT SIGNAL 1 32  
 FT CHAIN 33 769  
 FT DOMAIN 33 717 P-SELECTIN.  
 FT TRANSMEM 718 734 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 735 769 POTENTIAL.  
 FT DOMAIN 58 158 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 159 195 C-TYPE LECTIN.  
 FT DOMAIN 199 258 EGF-LIKE.  
 FT DOMAIN 261 320 SUSHI 1.  
 FT DOMAIN 323 382 SUSHI 2.  
 FT DOMAIN 385 444 SUSHI 3.  
 FT DOMAIN 447 506 SUSHI 4.  
 FT DOMAIN 509 568 SUSHI 5.  
 FT DOMAIN 580 639 SUSHI 6.  
 FT DOMAIN 642 701 SUSHI 7.  
 FT DOMAIN 701 769 SUSHI 8.  
 FT DISULFID 131 150 BY SIMILARITY.  
 FT DISULFID 163 174 BY SIMILARITY.  
 FT DISULFID 183 194 BY SIMILARITY.  
 FT DISULFID 185 194 BY SIMILARITY.  
 FT DISULFID 200 244 BY SIMILARITY.  
 FT DISULFID 230 257 BY SIMILARITY.  
 FT DISULFID 262 306 BY SIMILARITY.  
 FT DISULFID 319 368 BY SIMILARITY.  
 FT DISULFID 324 368 BY SIMILARITY.  
 FT DISULFID 354 381 BY SIMILARITY.  
 FT DISULFID 386 430 BY SIMILARITY.  
 FT DISULFID 416 443 BY SIMILARITY.  
 FT DISULFID 448 492 BY SIMILARITY.

FT DISULFID 478 505 BY SIMILARITY.  
 FT DISULFID 510 554 BY SIMILARITY.  
 FT DISULFID 540 567 BY SIMILARITY.  
 FT DISULFID 581 625 BY SIMILARITY.  
 FT DISULFID 611 638 BY SIMILARITY.  
 FT DISULFID 643 687 BY SIMILARITY.  
 FT DISULFID 673 700 BY SIMILARITY.  
 FT CARBOHYD 54 54 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 80 80 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 180 180 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 212 212 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 219 219 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 398 398 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 604 604 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 655 655 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 662 662 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 680 680 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 757 760 ENDOCYTOSIS SIGNAL (PROBABLE).  
 FT VARIANT 566 566 S -> T.  
 FT VARIANT 579 579 L -> V.  
 SQ SEQUENCE 769 AA; 84317 MW; 23E42575060FAB15 CRC64;  
 Query Match 40.9%; Score 300; DB 1; Length 769;  
 Best Local Similarity 39.8%; Pred. No. 6,9e-22;  
 Matches 49; Conservative 24; Mismatches 30; Indels 0; Gaps 0;  
 QY 1 CEPLAPELLGTMDCTHPNFSFSSQAFSCSGTMTNGIEFTGCPFGNWSPEPTCOV 60  
 DB 200 CSEFDPQVNHVNCSPRLNFSKSCSFHACAGVALNPREFELASGIWINSPPQVVA 259  
 QY 61 IOCEPLAPDLGIMNCSPHPLASFSTACTFICSGTELGKTKTCSSGIMNSPDC 120  
 DB 260 VCCPALKEPQSCSMSCFHSKAKAFQHSQSCSFCEGFTLVGEVVCALGWTATPVYC 319  
 QY 121 OKL 123  
 DB 320 KAI 322  
 RESULT 15  
 LEM2\_RAT  
 ID LEM2\_RAT STANDARD; PRT; 549 AA.  
 AC P98105;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
 DE (CD62E).  
 GN SELE OR ELAM-1.  
 GN Rattus norvegicus (Rat).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lung;  
 RA Rosenbloom C.L., Auchampach J.A., Anderson D.C., Manning A.M.;  
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.  
 CC -1 FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
 MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
 ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
 POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
 GLYCOPOLYIDS).  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L25527; AAA41113.1; -.  
DR HSSP: P15681; 1KJA.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00008; EGF; 1.  
DR Pfam: PF00059; lectin\_c; 1.  
DR Pfam: PF00084; sushi; 5.  
DR SMART: SM00032; CCP; 5.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00181; EGF; 1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS01186; EGF\_2; 1.  
DR PROSITE: PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE: PS50041; C-TYPE\_LECTIN\_2; 1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selection; Signal; Sushi; Repeat  
FT SIGNAL 1 21  
FT CHAIN 22 549  
FT DOMAIN 22 494  
FT TRANSMEM 495 516  
FT DOMAIN 517 549  
FT DOMAIN 38 138  
FT DOMAIN 139 175  
FT DOMAIN 179 239  
FT DOMAIN 242 301  
FT DOMAIN 304 364  
FT DOMAIN 367 427  
FT DOMAIN 430 486  
FT DISULFID 40 138  
FT DISULFID 111 130  
FT DISULFID 143 154  
FT DISULFID 148 163  
FT DISULFID 165 174  
FT DISULFID 180 225  
FT DISULFID 210 238  
FT DISULFID 243 287  
FT DISULFID 273 300  
FT DISULFID 305 350  
FT DISULFID 336 363  
FT DISULFID 368 413  
FT DISULFID 399 426  
FT DISULFID 431 472  
FT DISULFID 458 485  
FT CARBOHYD 25 25  
FT CARBOHYD 60 60  
FT CARBOHYD 145 145  
FT CARBOHYD 192 192  
FT CARBOHYD 203 203  
FT CARBOHYD 266 266  
FT CARBOHYD 313 313  
FT CARBOHYD 320 320  
FT CARBOHYD 333 333  
FT CARBOHYD 441 441  
FT CARBOHYD 465 465  
SQ SEQUENCE 549 AA; 60079 MW; 85CECBDB7B0144C8 CRC64;

Query Match 40.9%; Score 299.5; DB 1; Length 549;  
Best Local Similarity 41.0%; Pred. No. 5,5e-22;  
Matches 50; Conservative 23; Mismatches 48; Indels 1; Gaps 1;

OY 1 CEPLAEPLGMDCTHPEFGNFSFSCAFSCSGTNLGTEET-CGPFQNWSSPEPTCQ 59  
DB 180 CQDETFPHGSLNCTHPEFGFSTNNSGSCFSCERGIYVSSMETTVRCTSGSESNAPACH 239

OY 60 VICEPLSADPLGIMNCSHPLASFSTACTFCSGTELLGKKKTCSSGIMNSPST 119  
DB 240 VEEKALTPAHAGVGRKSSNNGSYPMWTTCTFDEEGYRRVGAQNLCTSSGWDNEKPS 239  
OY 120 CK 121  
DB 300 CK 301

RESULT 16  
LENN\_HUMAN  
ID LEEN2\_HUMAN STANDARD; PRT; 610 AA.  
AC P15681; P16111;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
DE (CD62E).  
GN SELE OR ELAM1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90175359; PubMed=1689848;  
RA Hession C., Osborn L., Goff D., Chl-Rosso G., Vassallo C.,  
RA Pasek M., Pittack C., Tizard R., Geelz S., McCarthy K., Hopple S.,  
RA Lobb R.,  
RT "Endothelial leukocyte adhesion molecule 1: direct expression cloning  
RT and functional interactions."  
RL Proc. Natl. Acad. Sci. U.S.A. 87:1673-1677(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89162047; PubMed=2466335;  
RA Bevilacqua M.P., Stengelin S., Gimbrone M.A. Jr., Seed B.,  
RT "Endothelial leukocyte adhesion molecule 1: an inducible receptor for  
RT neutrophils related to complement regulatory proteins and lectins."  
RL Science 243:1160-1165(1989).  
RN [3]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91115870; PubMed=1703529;  
RA Collins T., Williams A., Johnston G.I., Kim J., Eddy R., Shows T.,  
RA Gimbrone M.A. Jr., Bevilacqua M.P.,  
RT "Structure and chromosomal location of the gene for endothelial-  
RT leukocyte adhesion molecule 1."  
RL J. Biol. Chem. 266:2466-2473(1991).  
RN [4]  
RP LIGAND.  
RX MEDLINE=91068005; PubMed=1701274;  
RA Phillips M.L., Nudelman E., Gaeta F.C., Perez M., Singhal A.K.,  
RA Hakomori S., Paulson J.C.,  
RT "ELAM-1 mediates cell adhesion by recognition of a carbohydrate  
RT ligand, sialyl-Lex."  
RL Science 250:1130-1132(1990).  
RN [5]  
RP 3D-STRUCTURE MODELING OF LECTIN DOMAIN.  
RX MEDLINE=93202275; PubMed=7681016;  
RA Mills A.,  
RT "Modelling the carbohydrate recognition domain of human E-selectin."  
RL FEBS Lett. 319:5-11(1993).  
RN [6]  
RP X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 22-178.  
RX MEDLINE=94150646; PubMed=7509040;  
RA Graves B.J., Crowther R.L., Chandran C., Rumberger J.M., Li S.,  
RA Huang K.-S., Presky D.H., Familletti P.C., Wollitzky B.A., Burns D.K.,  
RT "Insight into E-selectin/ligand interaction from the crystal  
RT structure and mutagenesis of the Lec/EGF domains."  
RL Nature 367:532-538(1994).  
RN [7]  
RP VARIANT ARG-149.  
RX MEDLINE=95179107; PubMed=7533025;

RA Wenzel K., Felix S., Kleber F.X., Brachold R., Menke T., Schatke S.,  
RA Schulte K.L., Glaeser C., Rohde K., Baumann G., Speer A.;  
RA "E-selectin polymorphism and atherosclerosis: an association study.";  
RL Hum. Mol. Genet. 3:1935-1937(1994).  
RN [8]  
RP VARIANT ARG-149.  
RX MEDLINE-99134508; PubMed-9933738;  
RA Ye S.-Q., Usher D., Virgill D., Zhang L.-Q., Yochim S.E., Gupta R.;  
RT "A p61 polymorphism detects the mutation of serine-128 to arginine in  
RT CD 62E gene - a risk factor for coronary artery disease.";  
RL J. Biomed. Sci. 6:18-21(1999).  
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCCOSYLATED DERIVATIVES OF  
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
CC GLYCOPOLYIDS).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- POLYMORPHISM: A POLYMORPHISM IN POSITION 149 IS ASSOCIATED WITH A  
CC HIGHER RISK OF CORONARY ARTERY DISEASE (CAD). A SIGNIFICANTLY  
CC HIGHER MUTATION FREQUENCY (ARG-149) IS OBSERVED IN PATIENTS WITH  
CC ANGIOGRAPHICALLY PROVEN SEVERE ATHEROSCLEROSIS COMPARED WITH AN  
CC UNSELECTED POPULATION (SER-149).  
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMs.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.  
CC -1- DATABASE: NAME-PROW; NOTE-CD guide CD62E entry;  
CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd62e.htm".  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
DR EMBL: M30640; AAA52377.1; -  
DR EMBL: M61893; AAA52375.1; -  
DR EMBL: M61885; AAA52375.1; JOINED.  
DR EMBL: M61887; AAA52375.1; JOINED.  
DR EMBL: M61888; AAA52375.1; JOINED.  
DR EMBL: M61890; AAA52375.1; JOINED.  
DR EMBL: M61891; AAA52375.1; JOINED.  
DR EMBL: M61892; AAA52375.1; JOINED.  
DR EMBL: M24736; AAA52376.1; -  
DR PIR: A32606; A32606.  
DR PIR: A35046; A35046.  
DR PIR: A38615; A38615.  
DR PDB: 1ESL; 3I-AUG-94.  
DR PDB: 1KTA; 03-APR-96.  
DR MIM: 131210; -  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
DR InterPro: IPR001304; Lectin\_c.  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00059; Lectin\_c; 1.  
DR Pfam: PF00084; sushi; 6.  
DR PRINTS: PRO0343; SELECTIN.  
DR SMART: SM00032; CCP; 6.  
DR SMART: SM00034; CLECT; 1.  
DR SMART: SM00181; EGF\_1.  
DR PROSITE: PS00022; EGF\_1; 1.  
DR PROSITE: PS00186; EGF\_2; 1.  
DR PROSITE: PS00615; C\_TYPE\_LECTIN\_1; 1.  
DR PROSITE: PS00041; C\_TYPE\_LECTIN\_2; 1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat; Polymorphism; 3D-structure.  
FT SIGNAL 1 21  
FT CHAIN 1 610  
FT DOMAIN 22 556  
FT TRANSMEM 557 578  
FT E-SELECTIN.  
FT EXTRACELLULAR (POTENTIAL).  
FT POTENTIAL.

FT	DOMAIN	579	610	CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	38	138	C-TYPE LECTIN (SHORT FORM).
FT	DOMAIN	139	175	EGF-LIKE.
FT	DOMAIN	179	238	SUSHI 1.
FT	DOMAIN	241	300	SUSHI 2.
FT	DOMAIN	303	363	SUSHI 3.
FT	DOMAIN	366	426	SUSHI 4.
FT	DOMAIN	429	489	SUSHI 5.
FT	DOMAIN	492	548	SUSHI 6.
FT	DISULFID	40	138	
FT	DISULFID	111	130	
FT	DISULFID	143	154	
FT	DISULFID	148	163	
FT	DISULFID	165	174	
FT	DISULFID	180	224	BY SIMILARITY.
FT	DISULFID	210	237	BY SIMILARITY.
FT	DISULFID	242	286	BY SIMILARITY.
FT	DISULFID	272	299	BY SIMILARITY.
FT	DISULFID	304	349	BY SIMILARITY.
FT	DISULFID	335	362	BY SIMILARITY.
FT	DISULFID	367	412	BY SIMILARITY.
FT	DISULFID	398	425	BY SIMILARITY.
FT	DISULFID	430	475	BY SIMILARITY.
FT	DISULFID	461	488	BY SIMILARITY.
FT	DISULFID	493	534	BY SIMILARITY.
FT	DISULFID	520	547	BY SIMILARITY.
FT	CARBOHYD	25	25	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	145	145	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	160	160	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	179	179	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	199	199	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	203	203	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	265	265	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	312	312	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	332	332	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	503	503	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	527	527	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	VARIANT	130	130	C -> W (IN DBSNP:5360).
FT	VARIANT	149	149	/FTID=VAR_011790.
FT	VARIANT	149	149	S -> R (ASSOCIATED WITH A RISK FACTOR FOR CAD).
FT	VARIANT	295	295	/FTID=VAR_004191.
FT	VARIANT	295	295	E -> K (IN DBSNP:5364).
FT	VARIANT	421	421	/FTID=VAR_011791.
FT	VARIANT	421	421	E -> Q (IN DBSNP:5366).
FT	VARIANT	468	468	H -> Y (IN DBSNP:5368).
FT	VARIANT	575	575	/FTID=VAR_011793.
FT	VARIANT	575	575	L -> F (IN DBSNP:5355).
FT	VARIANT	575	575	/FTID=VAR_011794.
FT	SEQUENCE	610 AA; 66655 KM; 7D43E3C0D1229229 CRC64;		

Query Match 40.4%; Score 296; DB 1; Length 610;  
Best local Similarity 41.3%; Pred. No. 1,3e-21;  
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY	1	CEPLEAPELCTMDCTHPFGNFSSQCAFSCEGTNLGTIEETTCGPFGMSSPEPTCOV	60
DB	180	CTALESPEHGSILVCSHPLGNFSSCSISCDRGVLPSSMETQMCSGEMSAPIPCNV	239
QY	61	IOCEPLSAPDLGIMNCNSHPLASFSFTSACFTISEGELGKKTKTCESGGSINSPIC	120
DB	240	VECDAVTNPANGFVECPQNGSPFWNTCTFDEGEELMGASLOCTSSGNDNEKPTC	299
QY	121	Q 121	
DB	300	K 300	

RESULT 17  
LEM3\_BOVIN  
ID LEM3\_BOVIN STANDARD; PRT; 646 AA.

AC P42201;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE P-selectin precursor (granule membrane protein 140) (GMP-140) (PADGEM)  
 DE (CD62P) (leukocyte-endothelial cell adhesion molecule 3) (LECAM3).  
 GN SELEP.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Capillary endothelium;  
 RX MEDLINE=93249394; PubMed=7683458;  
 RA Strubel N.A., Nguyen M., Kansas G.S., Tedder T.F., Bischoff J.;  
 RT "Isolation and characterization of a bovine cDNA encoding a  
 RL functional homology of human P-selectin.";  
 CC Biochem. Biophys. Res. Commun. 192:338-344(1993).  
 CC -1 FUNCTION: CA(2+)-DEPENDENT RECEPTOR FOR MELOID CELLS THAT BINDS  
 CC TO CARBOHYDRATES ON NEUTROPHILS AND MONOCYTES. MEDIATES THE  
 CC INTERACTION OF ACTIVATED ENDOTHELIAL CELLS OR PLATELETS WITH  
 CC LEUKOCYTES. THE LIGAND RECOGNIZED IS STALBU-LEWIS X.  
 CC -1 SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1 TISSUE SPECIFICITY: STORED IN THE ALPHA-GRANULES OF PLATELETS  
 CC AND WEIBEL-PALADE BODIES OF ENDOTHELIAL CELLS. UPON CELL  
 CC ACTIVATION BY AGONISTS, P-SELECTIN IS TRANSPORTED RAPIDLY TO  
 CC THE CELL SURFACE.  
 CC -1 SIMILARITY: TO OTHER SELECTINS/LECAMs.  
 CC -1 SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
 CC -1 SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS; BOVINE P-LECTIN LACKS  
 CC THE HUMAN SUSHI-3, -4 AND -7 EQUIVALENTS.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC  
 CC EMBL: L12041; AAA30743.1; -  
 CC HSSP: P16109; IFSB.  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR002396; Selectin.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR InterPro: IPR001304; Lectin\_c.  
 DR Pfam: PF00008; EGF; 1.  
 DR Pfam: PF00059; lectin\_c; 1.  
 DR PRINTS: PR00343; SELECTIN.  
 DR SMART: SM00032; CCP; 6.  
 DR SMART: SM00034; CLECT; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR PROSITE: PS00022; EGF\_1; 1.  
 DR PROSITE: PS01186; EGF\_2; 1.  
 DR PROSITE: PS00615; C-TYPE LECTIN\_1; 1.  
 DR PROSITE: PS50041; C-TYPE LECTIN\_2; 1.  
 KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
 KW Selectin; Signal; Sush1; Repeat.  
 FT SIGNAL 1 41  
 FT CHAIN 42 646  
 FT DOMAIN 42 587  
 FT TRANSLEM 588 611  
 FT DOMAIN 612 646  
 FT DOMAIN 58 158  
 FT DOMAIN 159 195  
 FT DOMAIN 199 258  
 FT DOMAIN 261 320  
 FT DOMAIN 323 382  
 FT DOMAIN 385 444

FT DOMAIN 457 516  
 FT SUSHI 5.  
 FT DOMAIN 519 578  
 FT SUSHI 6.  
 FT DISULFID 60 158  
 FT BY SIMILARITY.  
 FT DISULFID 131 150  
 FT BY SIMILARITY.  
 FT DISULFID 163 174  
 FT BY SIMILARITY.  
 FT DISULFID 168 183  
 FT BY SIMILARITY.  
 FT DISULFID 185 194  
 FT BY SIMILARITY.  
 FT DISULFID 200 244  
 FT BY SIMILARITY.  
 FT DISULFID 230 257  
 FT BY SIMILARITY.  
 FT DISULFID 262 306  
 FT BY SIMILARITY.  
 FT DISULFID 292 319  
 FT BY SIMILARITY.  
 FT DISULFID 324 368  
 FT BY SIMILARITY.  
 FT DISULFID 354 381  
 FT BY SIMILARITY.  
 FT DISULFID 386 430  
 FT BY SIMILARITY.  
 FT DISULFID 416 443  
 FT BY SIMILARITY.  
 FT DISULFID 458 502  
 FT BY SIMILARITY.  
 FT DISULFID 488 515  
 FT BY SIMILARITY.  
 FT DISULFID 520 564  
 FT BY SIMILARITY.  
 FT DISULFID 550 577  
 FT BY SIMILARITY.  
 FT CARBOHYD 48 48  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 54 54  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 80 80  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 180 180  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 212 212  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 219 219  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 336 336  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 481 481  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 532 532  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 539 539  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 557 557  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SITE 634 637  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 646 AA; 71229 MW; 573912A4627A6ACA CRC64;  
 Query Match 39.68; Score 290; DB 1; Length 646;  
 Best Local Similarity 39.08; Pred. No. 5; Se-21;  
 Matches 48; Conservative 24; Mismatches 51; Indels 0; Gaps 0;  
 QY 1 CEPLEAPLGTMDCTHPGFNFSSOCAFSCSEGTNTGTIEETTCGPGFMWSSPEPTCOV 60  
 DB 200 CGEPDLQGHVIMNCSHPLGNFSFNSHCSFHCABEYALNGSELECLASGIWTNSPPCVA 259  
 QY 61 IQCEPLSAPDLGIMNCSHPLASFSTACTGCTEGTLEIGKKTKTCESSGIMNSPSPIC 120  
 DB 260 VQCPALKSPEGSMSCVQSAFAFQHSSCSFSCBEGALVGPVEVHTALGVMTAPTPVC 319  
 QY 121 QKL 123  
 DB 320 KAL 322  
 RESULT 18  
 LEM2\_RABIT  
 ID LEM2\_RABIT STANDARD; PRT; 551 AA.  
 AC P27113;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
 DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
 DE (CD62E).  
 GN SELE.  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OC NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Heart;  
 RX MEDLINE=92189729; PubMed=1372169;  
 RA Lariagan J.D., Tsang T.C., Rumberger J.M., Burns D.K.;  
 RT "Characterization of cDNA and genomic sequences encoding rabbit  
 RT ELAM-1: conservation of structure and functional interactions with



RT leukocytes.";  
RL DNA Cell Biol. 11:149-162(1992).  
CC -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
CC GLYCOLIPIDS).  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- INDUCTION: BY CYTOKINES.  
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 5 SUSHI (SCR) DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC  
CC EMBL: M91004; AAA31243.1; -.  
CC EMBL: M91005; AAA31244.1; -.  
CC HSSP: P16581; IKAJ.  
DR InterPro: IPR000561; EGF-like.  
DR InterPro: IPR002396; Selectin.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR InterPro: IPR001304; lectin\_c.  
DR Pfam: PF00008; EGF\_1.  
DR Pfam: PF00059; lectin\_c\_1.  
DR Pfam: PF00084; sush1\_5.  
DR PRINTS: PR00343; SELECTIN.  
DR SMART: SM00032; CCP\_5.  
DR SMART: SM00034; CLECT\_1.  
DR SMART: SM00181; EGF\_1.  
DR PROSITE: PS00022; EGF\_1\_1.  
DR PROSITE: PS01186; EGF\_2\_1.  
DR PROSITE: PS00615; C-TYPE\_LLECTIN\_1.  
DR PROSITE: PS50041; C-TYPE\_LLECTIN\_2\_1.  
KW Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sush1; Repeat.  
FT SIGNAL 1 23  
FT CHAIN 24 551 E-SELECTIN.  
FT DOMAIN 24 495 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 496 517 POTENTIAL.  
FT DOMAIN 518 551 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 40 140 C-TYPE LECTIN (SHORT FORM).  
FT DOMAIN 141 177 EGF-LIKE.  
FT DOMAIN 181 240 SUSHI 1.  
FT DOMAIN 243 302 SUSHI 2.  
FT DOMAIN 305 365 SUSHI 3.  
FT DOMAIN 368 428 SUSHI 4.  
FT DOMAIN 431 487 SUSHI 5.  
FT DISULFID 42 140 BY SIMILARITY.  
FT DISULFID 113 132 BY SIMILARITY.  
FT DISULFID 145 156 BY SIMILARITY.  
FT DISULFID 150 165 BY SIMILARITY.  
FT DISULFID 167 176 BY SIMILARITY.  
FT DISULFID 182 226 BY SIMILARITY.  
FT DISULFID 212 239 BY SIMILARITY.  
FT DISULFID 244 288 BY SIMILARITY.  
FT DISULFID 274 301 BY SIMILARITY.  
FT DISULFID 306 351 BY SIMILARITY.  
FT DISULFID 337 364 BY SIMILARITY.  
FT DISULFID 369 414 BY SIMILARITY.  
FT DISULFID 400 427 BY SIMILARITY.  
FT DISULFID 432 473 BY SIMILARITY.  
FT DISULFID 459 486 BY SIMILARITY.  
FT CARBOHYD 32 32 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 45 45 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 162 162 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 194 194 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 201 201 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 205 205 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 267 267 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 314 314 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 321 321 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 442 442 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 466 466 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 308 308 T -> A (IN REF. 1; AAA31244).  
FT CONFLICT 328 328 T -> A (IN REF. 1; AAA31244).  
FT CONFLICT 491 491 A -> V (IN REF. 1; AAA31244).  
SQ SEQUENCE 551 AA; 60346 MW; 23BC8A8B3B23240E CRC64;  
  
Query Match 39.3%; Score 288; DB 1; Length 551;  
Best Local Similarity 40.5%; Pred. No. 7,3e-21;  
Matches 49; Conservative 20; Mismatches 52; Indels 0; Gaps 0;  
  
QY 1 CEPLAEPETGMDCTHPFNFSFSSQCARSCSGTNLGTIEETGCGPFNMWSPEPTCOV 60  
DB 182 CEAQVOPQHGSLNCTHPFNFSFSSQCARSCSGTNLGTIEETGCGPFNMWSPEPTCOV 241  
QY 61 IOCEPLAPDLGIMNCSDPLASFTSACTFTCSGTELLGKRRKTCSSGIMNSPPTC 120  
DB 242 VECDFWGRKPRANGDVKCSFGSAPVMTCTCFDCEBGFLLGARSLOCTSSGSDNEKPTC 301  
QY 121 Q 121  
DB 302 K 302  
  
RESULT 19  
LEM2\_BOVIN  
ID P98107; STANDAR; PRT; 485 AA.  
AC 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
DE (ELAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
DE (CD62E).  
GN SELE.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Adrenal gland;  
RX MEDLINE=93382537; PubMed=7690465;  
RA Nguyen M., Strubel N.A., Bischoff J.;  
RT "A role for sialyl Lewis-X/A glycoconjugates in capillary  
morphogenesis.";  
RT Nature 365:267-269(1993).  
RL -1- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
CC ELAM-1 IS SIALYL-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
CC POLYLACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
CC GLYCOLIPIDS).  
CC -1- FUNCTION: INVOLVED IN CAPILLARY MORPHOGENESIS BY BINDING TO A  
CC BOVINE CAPILLARY ENDOTHELIAL (BCE) CELL. SIALYL LEWIS X AND/OR  
CC SIALYL LEWIS A-CONTAINING LIGAND (PROBABLY).  
CC -1- SIMILARITY: TO OTHER SELECTINS/LECAMS.  
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS; BOVINE E-LECTIN LACKS  
CC THE HUMAN SUSHI-4 AND 5 EQUIVALENTS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
DR EMBL; L12039; AAA02991.1; -.  
DR HSSP; P16581; 1ESL.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR000742; EGF 2.  
DR InterPro; IPR002396; Selectin.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR InterPro; IPR001304; lectin\_c.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi; 4.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 4.  
DR SMART; SM00034; CLECT; 1.  
DR SMART; SM00181; EGF; 1.  
DR PROSITE; PS00022; EGF\_1; 1.  
DR PROSITE; PS01186; EGF 2; 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_1; 1.  
DR PROSITE; PS00615; C-TYPE\_LECTIN\_2; 1.  
DR Cell adhesion; Transmembrane; Glycoprotein; EGF-like domain; Lectin;  
KW Selectin; Signal; Sushi; Repeat.  
FT SIGNAL 1 22  
FT CHAIN 1 22  
FT DOMAIN 23 485  
FT TRANSMEM 431 453  
FT DOMAIN 454 485  
FT DOMAIN 39 139  
FT DOMAIN 140 176  
FT DOMAIN 180 238  
FT DOMAIN 241 300  
FT DOMAIN 303 363  
FT DOMAIN 366 422  
FT DOMAIN 41 139  
FT DISULFID 112 131  
FT DISULFID 144 155  
FT DISULFID 149 164  
FT DISULFID 166 175  
FT DISULFID 181 224  
FT DISULFID 210 237  
FT DISULFID 242 286  
FT DISULFID 272 299  
FT DISULFID 304 349  
FT DISULFID 335 362  
FT DISULFID 367 408  
FT DISULFID 394 421  
FT CARBOHYD 61 61  
FT CARBOHYD 79 79  
FT CARBOHYD 88 88  
FT CARBOHYD 161 161  
FT CARBOHYD 203 203  
FT CARBOHYD 265 265  
FT CARBOHYD 312 312  
FT CARBOHYD 316 316  
FT CARBOHYD 379 379  
FT CARBOHYD 401 401  
SQ SEQUENCE 485 AA; 53200 MW; AE931C9B521E3904 CRC64;

Query Match 37.0%; Score 271.5; DB 1; Length 485;  
Best local Similarity 41.3%; Pred. No. 2,6e-19;  
Matches 50; Conservative 20; Mismatches 50; Indels 1; Gaps 1;

QY 1 CEPLAEELGTMDCTHPFGNFSFSSQCAFSCESTNLGTLEETTCGFGMSSPEPTQCV 60  
DB 181 CPAQKHHEGHILWC-NELGKFTYSSCSISCAEGYLPSSTREATRCMSSGEMSTPLPCNV 239  
QY 61 IOCEPLASAPDLGIMNCNHPPLASFSTFACFIQSEGTLELGGKKTIOESSGINSPEPIC 120  
DB 240 VKCDALSLNDNGVAVNCSPNHGSLPWNTTCTFECEQGYKLTGPQHLQCTSSGIMDNKQPTC 299

QY 121 0 121  
DB 300 K 300

RESULT 20  
ID LEM2\_PIG STANDARD; PRT; 484 AA.  
AC P98110;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 15-JUL-1998 (Rel. 36, Last annotation update)  
DE E-selectin precursor (Endothelial leukocyte adhesion molecule 1)  
DE (EIAM-1) (Leukocyte-endothelial cell adhesion molecule 2) (LECAM2)  
DE (CD62E).  
GN SELE.  
OS Sus scrofa (Pig).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
OX NCBI\_TaxID=9823;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Aortic endothelium;  
RX MEDLINE=95071392; PubMed=7526854;  
RA Rollins S.A., Evans M.J., Johnson K.K., Elliott E.A., Squinto S.P.,  
RA Matlis L.A., Rother R.P.;  
RT "Molecular and functional analysis of porcine E-selectin reveals a  
RT potential role in xenograft rejection."  
RL Biochem. Biophys. Res. Commun. 204:763-771(1994).  
[2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Aortic endothelium;  
RX MEDLINE=94271236; PubMed=7516159;  
RA Tsang Y.T.M., Haskard D.O., Robinson M.K.;  
RT "Cloning and expression kinetics of porcine vascular cell adhesion  
RT molecule."  
RL Biochem. Biophys. Res. Commun. 201:805-805(1994).  
CC -I- FUNCTION: EXPRESSED ON CYTOKINE INDUCED ENDOTHELIAL CELLS AND  
CC MEDIATES THEIR BINDING TO LEUKOCYTES. THE LIGAND RECOGNIZED BY  
CC ELAM-1 IS STALK-LEWIS X (ALPHA(1->3)FUCOSYLATED DERIVATIVES OF  
CC POLYACTOSAMINE THAT ARE FOUND AT THE NONREDUCING TERMINI OF  
CC GLYCOLIPIDS).  
CC -I- FUNCTION: PLAYS AN IMPORTANT ROLE IN ACUTE CELLULAR ALLOGRAFT  
CC REJECTION AND PROBABLY ALSO IN XENOGRAFT REJECTION.  
CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -I- SIMILARITY: TO OTHER SELECTINS/LECAMS.  
CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.  
CC -I- SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.  
CC -I- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS; PORCINE E-LECTIN LACKS  
CC THE HUMAN SUSHI-1 AND -4 EQUIVALENTS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L139076; AAA61545.1; -.  
DR EMBL; U08350; AAA21541.1; -.  
DR HSSP; P16581; 1ESL.  
DR InterPro; IPR000561; EGF-like.  
DR InterPro; IPR002396; Selectin.  
DR InterPro; IPR000436; Sushi\_SCR\_CCP.  
DR InterPro; IPR001304; lectin\_c.  
DR Pfam; PF00008; EGF; 1.  
DR Pfam; PF00059; lectin\_c; 1.  
DR Pfam; PF00084; sushi; 4.  
DR PRINTS; PR00343; SELECTIN.  
DR SMART; SM00032; CCP; 4.  
DR SMART; SM00034; CLECT; 1.





FT	SIGNAL	1	18	
FT	CHAIN	19	1231	COMPLEMENT FACTOR H.
FT	DOMAIN	20	81	SUSHI 1.
FT	DOMAIN	84	142	SUSHI 2.
FT	DOMAIN	145	206	SUSHI 3.
FT	DOMAIN	209	263	SUSHI 4.
FT	DOMAIN	266	321	SUSHI 5.
FT	DOMAIN	324	386	SUSHI 6.
FT	DOMAIN	388	443	SUSHI 7.
FT	DOMAIN	447	506	SUSHI 8.
FT	DOMAIN	508	565	SUSHI 9.
FT	DOMAIN	568	624	SUSHI 10.
FT	DOMAIN	629	685	SUSHI 11.
FT	DOMAIN	690	745	SUSHI 12.
FT	DOMAIN	752	804	SUSHI 13.
FT	DOMAIN	810	865	SUSHI 14.
FT	DOMAIN	869	927	SUSHI 15.
FT	DOMAIN	930	985	SUSHI 16.
FT	DOMAIN	988	1044	SUSHI 17.
FT	DOMAIN	1047	1103	SUSHI 18.
FT	DOMAIN	1108	1164	SUSHI 19.
FT	DOMAIN	1166	1229	SUSHI 20.
FT	DISULFID	21	66	BY SIMILARITY.
FT	DISULFID	52	80	BY SIMILARITY.
FT	DISULFID	85	129	BY SIMILARITY.
FT	DISULFID	114	141	BY SIMILARITY.
FT	DISULFID	146	192	BY SIMILARITY.
FT	DISULFID	178	205	BY SIMILARITY.
FT	DISULFID	210	251	BY SIMILARITY.
FT	DISULFID	237	262	BY SIMILARITY.
FT	DISULFID	267	309	BY SIMILARITY.
FT	DISULFID	294	320	BY SIMILARITY.
FT	DISULFID	325	374	BY SIMILARITY.
FT	DISULFID	357	385	BY SIMILARITY.
FT	DISULFID	389	431	BY SIMILARITY.
FT	DISULFID	416	442	BY SIMILARITY.
FT	DISULFID	448	494	BY SIMILARITY.
FT	DISULFID	477	505	BY SIMILARITY.
FT	DISULFID	509	553	BY SIMILARITY.
FT	DISULFID	536	564	BY SIMILARITY.
FT	DISULFID	569	611	BY SIMILARITY.
FT	DISULFID	597	623	BY SIMILARITY.
FT	DISULFID	630	673	BY SIMILARITY.
FT	DISULFID	659	684	BY SIMILARITY.
FT	DISULFID	691	733	BY SIMILARITY.
FT	DISULFID	719	744	BY SIMILARITY.
FT	DISULFID	753	792	BY SIMILARITY.
FT	DISULFID	781	803	BY SIMILARITY.
FT	DISULFID	811	853	BY SIMILARITY.
FT	DISULFID	839	864	BY SIMILARITY.
FT	DISULFID	870	915	BY SIMILARITY.
FT	DISULFID	901	926	BY SIMILARITY.
FT	DISULFID	931	973	BY SIMILARITY.
FT	DISULFID	959	984	BY SIMILARITY.
FT	DISULFID	989	1032	BY SIMILARITY.
FT	DISULFID	1018	1043	BY SIMILARITY.
FT	DISULFID	1047	1091	BY SIMILARITY.
FT	DISULFID	1077	1102	BY SIMILARITY.
FT	DISULFID	1109	1152	BY SIMILARITY.
FT	DISULFID	1138	1163	BY SIMILARITY.
FT	DISULFID	1167	1218	BY SIMILARITY.
FT	DISULFID	1201	1228	BY SIMILARITY.
FT	CARBOHYD	529	529	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	718	718	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	802	802	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	822	822	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	882	882	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	911	911	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	1029	1029	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	CARBOHYD	1095	1095	N-LINKED (GLCNAC. . .) (POTENTIAL)
FT	VARSPLIC	446	449	KTCS -> SFTL (IN ISOFORM 2).
FT	VARSPLIC	450	450	MISSING (IN ISOFORM 2).
FT	VARIANT	402	402	H -> Y.

			/FTID=VAR_001979.
FF	CONFLICT	21	C -> Q (IN REF. 3).
FF	CONFLICT	30	T -> V (IN REF. 3).
FF	CONFLICT	34	T -> Q (IN REF. 3).
FF	CONFLICT	53	RP -> IL (IN REF. 2).
FF	STRAND	870	
FF	STRAND	876	

Query Match                  23.6%; Score 173; DB 1; Length 1231;  
Best Local Similarity      28.8%; Pred. No. 2,7e-09;

Matches	36:	Conservative	25:	Mismatches	48:	Indels	16:	Gaps
---------	-----	--------------	-----	------------	-----	--------	-----	------

  

```

Oy      14 CTH-----PFGNFS-----FSSQCAFCSGSEGTNLTG-IETTGGPFGNMSSPEPTCOV   60
          | | | | : 
Db       85 CGHPEDFPFGFFTLTGVVEFYGVKAVYTCNEGVLGLGEINTRECDTDG-WINDIPICEV    143
Oy      61 IQCEPLSPDGC--IMNCSHPIASFSTSACTICTSGSRTLLICKKTITLESSGISNSNP     118
          ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      144 VKCLEVTAPENGAKTVSSAMEPRDEHYHGAVFRCVNGSYKIIEGDDEMHSCDDGFWSKEKP   203
Oy      119 ICGKL 123
          ||:
Db      204 KCVET 208

```

```

RESULT  23
CRL_HUMAN ID CRL_HUMAN STANDARD; PRT; 2039 AA. ~
AC        PI7927;
DT       01-NOV-1990 (Rel. 16, Created)
DT       01-NOV-1990 (Rel. 16, Last sequence update)
DT       16-OCT-2001 (Rel. 40, Last annotation update)
DE       Complement receptor type 1 precursor (C3b/C4b receptor) (CD35 antigen).
EN
GN       CRI OR C3BR.
OS       Homo sapiens (Human).
OC       Eumaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX       NCBI_TaxId:9606;
RN       [1]
RP       SEQUENCE FROM N.A. MEDLINE=89035992; PubMed=2972794; RX       RA       Klickstein L.B., Bartow T.J., Miletic V., Rabson L.D., Smith J.A., Fearon D.T.; "Identification of distinct C3b and C4b recognition sites in the human C3b/C4b receptor (CRI, CD35) by deletion mutagenesis.", J. Exp. Med. 166:1699-1717(1988).
RN       [2]
RP       SEQUENCE OF 503-2039 FROM N.A. MEDLINE=87168191; PubMed=2951479; RX       RA       Klickstein L.B., Wong W.W., Smith J.A., Weis J.H., Wilson J.G., Fearon D.T.; "Human C3b/C4b receptor (CRI). Demonstration of long homologous repeating domains that are composed of the short consensus repeats characteristics of C3/c4 binding proteins." J. Exp. Med. 165:1095-1112(1987).
RN       [3]
RP       SEQUENCE OF 761-783; 831-845 AND 1179-1195 FROM N.A. MEDLINE=86067975; PubMed=2933745; RX       RA       Wong W.W., Klickstein L.B., Smith J.A., Weis J.H., Fearon D.T.; "Identification of a partial cDNA clone for the human receptor for complement fragments C3b/C4b." Proc. Natl. Acad. Sci. U.S.A. 82:7711-7715(1985).
CC       CC         -I- FUNCTION: CRI, PRESENT ON ERYTHROCYTES, LEUKOCYTES, GLOMERULAR PODOCYTES, AND SPLENIC FOLLICULAR DENDRITIC CELLS. MEDIATES THE BINDING BY THESE CELLS OF PARTICLES AND IMMUNE COMPLEXES THAT HAVE ACTIVATED COMPLEMENT.
CC       CC         -I- SUBUNIT: MONOMER.
CC       CC         -I- SUBCELLULAR LOCATION: Type I membrane protein.
CC       CC         -I- POLYMORPHISM: CRI IS RESPONSIBLE FOR THE KNOPS BLOOD GROUP SYSTEM.
CC       CC         -I- MISCELLANEOUS: SEVEN SHORT CONSENSUS REPEATS (SCR) CONSTITUTE A LONG HOMOLOGOUS REPEAT (LHR), THE N-TERMINAL TWO SRCS OF LHR-A CONTAINED A SITE DETERRMINING CA SPECIFICITY. AND THE N-TERMINAL

```

CC	TWO SCRS OF LHR-BAND -C EACH HAD A SITE DETERMINING C3 SPECIFICITY.			FT	DISULFID	266	293	BY SIMILARITY.
CC	-1- MISCELLANEOUS: THIS IS THE SEQUENCE OF THE F ALLOTYPIC OF CRL.			FT	DISULFID	297	340	BY SIMILARITY.
CC	-1- SIMILARITY: CONTAINS 30 SUSHI (SCR) DOMAINS.			FT	DISULFID	326	353	BY SIMILARITY.
CC	-----			FT	DISULFID	358	400	BY SIMILARITY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			FT	DISULFID	421	470	BY SIMILARITY.
CC	-----			FT	DISULFID	450	487	BY SIMILARITY.
CC				FT	DISULFID	493	536	BY SIMILARITY.
CC				FT	DISULFID	523	549	BY SIMILARITY.
CC				FT	DISULFID	554	595	BY SIMILARITY.
CC				FT	DISULFID	581	611	BY SIMILARITY.
CC				FT	DISULFID	616	665	BY SIMILARITY.
DR	EMBL; M11569; AAA52297.1; -		FT	DISULFID	645	682	BY SIMILARITY.	
DR	EMBL; M11617; AAA52298.1; -		FT	DISULFID	688	730	BY SIMILARITY.	
DR	EMBL; M11618; AAA52299.1; -		FT	DISULFID	716	743	BY SIMILARITY.	
DR	EMBL; Y00816; CAA68755.1; -		FT	DISULFID	747	790	BY SIMILARITY.	
DR	EMBL; X05309; CAA28933.1; -		FT	DISULFID	776	803	BY SIMILARITY.	
DR	PIR; A28507; A28507.		FT	DISULFID	808	850	BY SIMILARITY.	
DR	PIR; A24748; A24748.		FT	DISULFID	836	866	BY SIMILARITY.	
DR	PIR; B24748; B24748.		FT	DISULFID	871	920	BY SIMILARITY.	
DR	PIR; C24748; C24748.		FT	DISULFID	900	937	BY SIMILARITY.	
DR	PIR; S03843; S03843.		FT	DISULFID	943	986	BY SIMILARITY.	
DR	HSSP; P08603; IHT.		FT	DISULFID	973	999	BY SIMILARITY.	
DR	MIM; 120620; -		FT	DISULFID	1004	1045	BY SIMILARITY.	
DR	InterPro: IPRO00436; Sushi_SCR_CCP.		FT	DISULFID	1031	1061	BY SIMILARITY.	
DR	Pfam: PF00084; sushi; 30.		FT	DISULFID	1066	1115	BY SIMILARITY.	
DR	SMART: SM00032; CCP; 30.		FT	DISULFID	1095	1132	BY SIMILARITY.	
KW	Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal; Receptor; Sushi; Blood group antigen.		FT	DISULFID	1138	1180	BY SIMILARITY.	
KW	Signal	1	41	FT	DISULFID	1166	1193	BY SIMILARITY.
FT	CHAIN	42	2039	FT	DISULFID	1197	1240	BY SIMILARITY.
FT	DOMAIN	42	1971	FT	DISULFID	1226	1253	BY SIMILARITY.
FT	TRANSMEM	1972	1996	FT	DISULFID	1258	1300	BY SIMILARITY.
FT	DOMAIN	1997	2039	FT	DISULFID	1316	1316	BY SIMILARITY.
FT	MOD_RES	42	42	FT	DISULFID	1321	1370	BY SIMILARITY.
FT	DOMAIN	42	100	FT	DISULFID	1350	1387	BY SIMILARITY.
FT	DOMAIN	103	162	FT	DISULFID	1396	1439	BY SIMILARITY.
FT	DOMAIN	165	233	FT	DISULFID	1426	1452	BY SIMILARITY.
FT	DOMAIN	237	294	FT	DISULFID	1457	1498	BY SIMILARITY.
FT	DOMAIN	296	354	FT	DISULFID	1484	1514	BY SIMILARITY.
FT	DOMAIN	357	417	FT	DISULFID	1519	1568	BY SIMILARITY.
FT	DOMAIN	420	488	FT	DISULFID	1548	1585	BY SIMILARITY.
FT	DOMAIN	492	550	FT	DISULFID	1591	1633	BY SIMILARITY.
FT	DOMAIN	553	612	FT	DISULFID	1619	1646	BY SIMILARITY.
FT	DOMAIN	615	683	FT	DISULFID	1650	1693	BY SIMILARITY.
FT	DOMAIN	687	744	FT	DISULFID	1679	1706	BY SIMILARITY.
FT	DOMAIN	746	804	FT	DISULFID	1711	1753	BY SIMILARITY.
FT	DOMAIN	807	867	FT	DISULFID	1739	1769	BY SIMILARITY.
FT	DOMAIN	870	938	FT	DISULFID	1774	1823	BY SIMILARITY.
FT	DOMAIN	942	1000	FT	DISULFID	1803	1840	BY SIMILARITY.
FT	DOMAIN	1003	1062	FT	DISULFID	1848	1891	BY SIMILARITY.
FT	DOMAIN	1065	1133	FT	DISULFID	1877	1904	BY SIMILARITY.
FT	DOMAIN	1137	1194	FT	DISULFID	1909	1952	BY SIMILARITY.
FT	DOMAIN	1196	1254	FT	DISULFID	1938	1965	BY SIMILARITY.
FT	DOMAIN	1257	1317	FT	CARBOHYD	56	56	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1320	1388	FT	CARBOHYD	252	252	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1395	1453	FT	CARBOHYD	410	410	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1456	1515	FT	CARBOHYD	447	447	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1518	1586	FT	CARBOHYD	509	509	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1590	1647	FT	CARBOHYD	578	578	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1649	1707	FT	CARBOHYD	702	702	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1710	1770	FT	CARBOHYD	860	860	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1773	1841	FT	CARBOHYD	897	897	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1847	1905	FT	CARBOHYD	959	959	N-LINKED (GLCNAC. . .)
FT	DOMAIN	1908	1966	FT	CARBOHYD	1028	1028	N-LINKED (GLCNAC. . .)
FT	DISULFID	43	86	FT	CARBOHYD	1152	1152	N-LINKED (GLCNAC. . .)
FT	DISULFID	73	99	FT	CARBOHYD	1310	1310	N-LINKED (GLCNAC. . .)
FT	DISULFID	104	145	FT	CARBOHYD	1481	1481	N-LINKED (GLCNAC. . .)
FT	DISULFID	131	161	FT	CARBOHYD	1504	1504	N-LINKED (GLCNAC. . .)
FT	DISULFID	166	215	FT	CARBOHYD	1534	1534	N-LINKED (GLCNAC. . .)
FT	DISULFID	195	232	FT	CARBOHYD	1540	1540	N-LINKED (GLCNAC. . .)
FT	DISULFID	238	280	FT	CARBOHYD	1605	1605	N-LINKED (GLCNAC. . .)
FT	DISULFID			FT	CARBOHYD	1763	1763	N-LINKED (GLCNAC. . .)
FT	DISULFID			FT	CARBOHYD	1908	1908	N-LINKED (GLCNAC. . .)



```

FT CARBOHYD 721 721 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 773 773 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 801 801 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1030 1030 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1061 1061 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1225 1225 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1234 AA, 139082 MW, C5AC02F341B957F7 CRC64;

Query Match 21.8%; Score 160; DB 1; Length 1234;
Best Local Similarity 28.0%; Pred. No. 5,1e-08;
Matches 35; Conservative 25; Mismatches 49; Indels 16; Gaps 5;

QY 14 CTH-----PFGN-----FSSQCAFSCSEGNLTG-IEETTCGPGNMSSPEPTQCV 60
DB 85 CGHPDGFPGFGRFLAVSGSFERGAKVYTCDDGYQLLGEIDYRECGADG-WINDIPICEV 143
QY 61 IOCEPLSAPDLG--IMNSHPLASFSTACTFICSECTELIGKTKTCSSGIMNSPSP 118
DB 144 VKCLPVTLENGRIYSGAETDQYVRFEGVRFECNSGKIEGHEITHSENGIMSNKP 203
QY 119 ICORL 123
DB 204 RCVEI 208

RESULT 25
DAFL_MOUSE STANDARD; PRT; 390 AA.
ID DAFL_MOUSE
AC 061475; 061397; P97732;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Complement decay-accelerating factor, GPI-anchored precursor
  (DAF-GPI).
GN DAF1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=95403982; Pubmed=7545711;
RA Spicer A.P., Seldin M.F., Gendler S.J.;
RT "Molecular cloning and chromosomal localization of the mouse decay-
  accelerating factor genes. Duplicated genes encode
  glycosylphosphatidylinositol-anchored and transmembrane forms.";
RL J. Immunol. 155:3079-3091(1995).
RN [2]
RP SEQUENCE OF 7-390 FROM N.A.
RC STRAIN=BALB/C; TISSUE=Spleen;
RX MEDLINE=96362213; Pubmed=8671624;
RA Fukunaka Y., Yasui A., Okada H.;
RT "Molecular cloning of murine decay accelerating factor by
  immunoscreening.";
RL Int. Immunol. 8:379-385(1996).
CC -1- FUNCTION: PROTECTION OF CELLS FROM COMPLEMENT-MEDIATED DAMAGE (BY
  SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: BRAIN, SECRETORY EPITHELIA, SKELETAL MUSCLE,
  LIVER, TESTES, THYMUS, SPLEEN AND LYMPH NODE.
CC -1- DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR
  FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
  ACTIVE SITE ON SCR3 (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
  (RCA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; L41366; AAB00091.1; -.
DR EMBL; D63679; BAA09830.1; -.
DR HSSP; P08603; IHCC.
DR MGD; MGI:104850; DAF1.
DR InterPro; IPR000436; Sushi_SCR_CCP.
DR Pfam; PF00064; sushi; 4.
DR SMART; SM00032; CCP; 4.
KW Complement pathway; Glycoprotein; Membrane; Repeat; GPI-anchor;
KW Signal; Sushi.
FT FT 1 34
FT CHAIN 35 361
FT FT 35 361
FT PROPEP 362 390
FT DOMAIN 35 95
FT DOMAIN 97 159
FT DOMAIN 162 221
FT DOMAIN 224 285
FT DOMAIN 288 364
FT DISULFID 65 94
FT DISULFID 98 145
FT DISULFID 129 158
FT DISULFID 163 204
FT DISULFID 190 220
FT DISULFID 225 267
FT CARBOHYD 187 187
FT CARBOHYD 262 262
FT LIPID 361 361
FT CONFLICT 9 7
FT CONFLICT 9 9
FT CONFLICT 91 83
FT CONFLICT 91 91
FT CONFLICT 135 135
FT CONFLICT 173 173
FT CONFLICT 180 180
SQ SEQUENCE 390 AA; 42618 MW; 4418721DFF47F8E7 CRC64;

Query Match 21.2%; Score 155.5; DB 1; Length 390;
Best Local Similarity 30.9%; Pred. No. 4,2e-08;
Matches 42; Conservative 16; Mismatches 63; Indels 15; Gaps 6;

QY 1 CEPLAPELGTMCCTH---PFGNFSSSQCAFSCSGTNLTGIEETTCGPGN---WSSP 54
DB 158 CKKSCPNPKDLNGHINIPGT-ILFGSEINFCNPGYRLVGVSSTFCSTGNTYDWDDE 216
QY 55 EPTQVIOCEPLSAPDLGIMNSHPLASFSTACTFICSECTELIGKTKTCSSGIMNSPSP 110
DB 217 FVPTCTHICPEPPKXINGIMRGESD--SYTISQVYIYSCDKGFLVGNASTICTYKSDV 274
QY 111 GIMNSPSPICQKIDS 126
DB 275 GQWSSPPPRC--IEKS 288

RESULT 26
LFC_TACTR
ID LFC_TACTR STANDARD; PRT; 1019 AA.
AC P28175;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Limulus clotting factor C precursor (EC 3.4.21.84) (FC).
OS Tachyplesus tridentatus (Japanese horseshoe crab).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;
OC Limulidae; Tachyplesus.
OX NCBI_TaxID=6853;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.

```



FT	DOMAIN	200	254	SUSH 2.
FT <td>DOMAIN</td> <td>260</td> <td>321</td> <td>SUSH 3.</td>	DOMAIN	260	321	SUSH 3.
FT <td>DOMAIN</td> <td>325</td> <td>421</td> <td>LCCL.</td>	DOMAIN	325	421	LCCL.
FT <td>DOMAIN</td> <td>436</td> <td>568</td> <td>C-TYPE LECTIN.</td>	DOMAIN	436	568	C-TYPE LECTIN.
FT <td>DOMAIN</td> <td>576</td> <td>634</td> <td>SUSH 4.</td>	DOMAIN	576	634	SUSH 4.
FT <td>DOMAIN</td> <td>685</td> <td>748</td> <td>SUSH 5.</td>	DOMAIN	685	748	SUSH 5.
FT <td>DOMAIN</td> <td>763</td> <td>1019</td> <td>SERINE PROTEASE.</td>	DOMAIN	763	1019	SERINE PROTEASE.
FT <td>ACT_SITE</td> <td>809</td> <td>809</td> <td>CHANGE RELAY SYSTEM (BY SIMILARITY).</td>	ACT_SITE	809	809	CHANGE RELAY SYSTEM (BY SIMILARITY).
FT <td>ACT_SITE</td> <td>865</td> <td>865</td> <td>CHANGE RELAY SYSTEM (BY SIMILARITY).</td>	ACT_SITE	865	865	CHANGE RELAY SYSTEM (BY SIMILARITY).
FT <td>ACT_SITE</td> <td>966</td> <td>966</td> <td>CHANGE RELAY SYSTEM (BY SIMILARITY).</td>	ACT_SITE	966	966	CHANGE RELAY SYSTEM (BY SIMILARITY).
FT <td>BINDING</td> <td>960</td> <td>960</td> <td>SUBSTRATE (BY SIMILARITY).</td>	BINDING	960	960	SUBSTRATE (BY SIMILARITY).
FT <td>DOMAIN</td> <td>643</td> <td>689</td> <td>PRO-RICH.</td>	DOMAIN	643	689	PRO-RICH.
FT <td>DISULFID</td> <td>106</td> <td>118</td> <td>BY SIMILARITY.</td>	DISULFID	106	118	BY SIMILARITY.
FT <td>DISULFID</td> <td>112</td> <td>126</td> <td>BY SIMILARITY.</td>	DISULFID	112	126	BY SIMILARITY.
FT <td>DISULFID</td> <td>127</td> <td>136</td> <td>BY SIMILARITY.</td>	DISULFID	127	136	BY SIMILARITY.
FT <td>DISULFID</td> <td>436</td> <td>447</td> <td>BY SIMILARITY.</td>	DISULFID	436	447	BY SIMILARITY.
FT <td>DISULFID</td> <td>464</td> <td>564</td> <td>BY SIMILARITY.</td>	DISULFID	464	564	BY SIMILARITY.
FT <td>DISULFID</td> <td>538</td> <td>556</td> <td>BY SIMILARITY.</td>	DISULFID	538	556	BY SIMILARITY.
FT <td>DISULFID</td> <td>794</td> <td>810</td> <td>BY SIMILARITY.</td>	DISULFID	794	810	BY SIMILARITY.
FT <td>DISULFID</td> <td>932</td> <td>951</td> <td>BY SIMILARITY.</td>	DISULFID	932	951	BY SIMILARITY.
FT <td>DISULFID</td> <td>962</td> <td>996</td> <td>BY SIMILARITY.</td>	DISULFID	962	996	BY SIMILARITY.
FT <td>CARBOHYD</td> <td>523</td> <td>523</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	523	523	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>534</td> <td>534</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	534	534	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>624</td> <td>624</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	624	624	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>740</td> <td>740</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	740	740	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>767</td> <td>767</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	767	767	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>CARBOHYD</td> <td>912</td> <td>912</td> <td>N-LINKED (GLCNAC. . .) (POTENTIAL).</td>	CARBOHYD	912	912	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT <td>VARSPLIC</td> <td>492</td> <td>498</td> <td>LTWIG -&gt; TDWYAT (IN SHORT ISOFORM).</td>	VARSPLIC	492	498	LTWIG -> TDWYAT (IN SHORT ISOFORM).
FT <td>VARSPLIC</td> <td>499</td> <td>1019</td> <td>MISSING (IN SHORT ISOFORM).</td>	VARSPLIC	499	1019	MISSING (IN SHORT ISOFORM).
SO <td>SEQUENCE</td> <td>1019 AA;</td> <td>112346 MW;</td> <td>5BC2864C6715289B CRC64;</td>	SEQUENCE	1019 AA;	112346 MW;	5BC2864C6715289B CRC64;
Query Match				
Best Local Similarity		21.2%;	Score 155.5;	DB 1; Length 1019;
Matches 33; Conservative		34.7%;	Pred. No. 1,1e-07;	
		16;	Mismatches 43;	Indels 3; Gaps 2;
OY	29 FSCSEGTMTJGIEETTCGPGFNWSSPEPTCOYIOCEPLSADPLGIMNCSHPLASFSTSA 88			
DB	166 YSCSFGFLKGVARISCPCNGMSFPPKRC-IRECAKYSSPEHGKVNA--PSGMIEGAT 222			
OY	89 CTFCSGELTGKKTICGSSGIMSNSPITCKL 123			
DB	223 LRFSCDSPYLLIGQETLTCQNGMGWSGQIPCKKL 257			
RESULT 27				
CR2_MOUSE	CR2_MOUSE	STANDARD:	PRT:	1025 AA.
AC	P19070:			
DT	01-NOV-1990 (Rel. 16, Created)			
DT	01-NOV-1990 (Rel. 16, Last sequence update)			
DT	01-FEB-1996 (Rel. 33, Last annotation update)			
DE	Complement receptor type 2 precursor (Cr2) (Complement C3d receptor).			
SN	CR2.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RM	[1]			
RM	SEQUENCE FROM N.A.			
RP	STRAIN=BALB/C;			
RX	MEDLINE=90229735; PubMed=2139457;			
RA	Fingeroth J.D.;			
RT	"Comparative structure and evolution of murine CR2. The homolog of			
RL	J. Immunol. 144:3458-3467(1990)."			
RM	[2]			
RM	SEQUENCE OF 12-1025 FROM N.A.			
RX	MEDLINE=91010789; PubMed=2145366;			
RA	Molina H., Kinoshita T., Inoue K., Carel J.C., Holers V.M.;			
RT	"A molecular and immunochemical characterization of mouse CR2.			
RT	Evidence for a single gene model of mouse complement receptors 1 and			
RT	2.";			









```

FT DISULFID 561 561 INTERCHAIN (WITH BETA CHAIN) (POTENTIAL).
FT CARBOHYD 66 66 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 221 221 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 525 525 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 602 602 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 610 AA: 68886 MW: D806B270E8A06B58 CRC64:

Query Match 19.8%; Score 145; DB 1; Length 610;
Best Local Similarity 30.7%; Pred. No. 7.1e-07;
Matches 35; Conservative 18; Mismatches 47; Indels 14; Gaps 4;

OY 20 NFSFSCAFSCSEGTNLGIEETTCGPPGN---WSPEPTCQVIOCEPLAPDGIIMNC 76
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 128 DYSFGESEIEFSCSEGYVIGSANSYCOLDKGVWSDPLPQCIIAKCEPPPT---ISNG 183
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 77 SHPLAS---FSPTACFTICSGTELGKKTIC---ESSGIWNSPPIQOKL 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 184 RHNGDEDEPTYYGSSVTTSCDRDFSLGKASISCKVENKTTIGWSPSPCKKV 237
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 32
C4BP_MOUSE STANDARD: PRT: 469 AA.
ID C4BP_MOUSE
AC P06607:
DT 01-AUG-1988 (Rel. 08, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE C4b-binding protein precursor (c4bp).
GN C4BP OR C4BP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88024997; PubMed=3663616.
RA Kristensen T., Ogata R.T., Chung L.P., Reid K.B.M., Tack B.F.;
RT "cDNA structure of murine C4b-binding protein, a regulatory component
  of the serum complement system."
RL Biochemistry 26:4668-4674(1987).
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
  ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
  (C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
  ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
  CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. ALPHA
  CHAIN BINDS C4b. IT INTERACTS ALSO WITH SERUM AMYLOID P COMPONENT.
CC -1- SUBUNIT: HOMOPENTAMER, NOT COVALENTLY LINKED. MOUSE LACKS THE
  BETA CHAIN OF C4BP.
CC -1- SIMILARITY: TO C4BP BETA CHAIN AND TO PIG APOLIPOPROTEIN R.
CC -1- SIMILARITY: CONTAINS 6 SUSHI (SCR) DOMAINS.
CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-44 IS THE INITIATOR.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
  between the Swiss Institute of Bioinformatics and the EMBL Outstation
  at the European Bioinformatics Institute. There are no restrictions on its
  use by non-profit institutions as long as its content is in no way
  modified and this statement is not removed. Usage by and for commercial
  entities requires a license agreement (See http://www.isb-sib.ch/announce/
  or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M17122; AAA37312.1; ALT_INIT.
DR PIR: A27117; NEMSC4.
DR HSSP: P10996; IVD.
DR MGD: MGI:88229; C4BP.
DR InterPro: IPR000436; Sush1_SCR_CCP.
DR Pfam: PF00084; sush1: 6.
DR SMART: SM00032; CCP; 6.
KM Complement pathway; Plasma; Glycoprotein; Repeat; Sush1; Signal.
FT SIGNAL 1 56
FT CHAIN 57 469 C4B-BINDING PROTEIN.
FT DOMAIN 57 116 SUSHI 1.
FT DOMAIN 119 177 SUSHI 2.

```

```

FT DOMAIN 180 241 SUSHI 3.
FT DOMAIN 244 300 SUSHI 4.
FT DOMAIN 302 356 SUSHI 5.
FT DOMAIN 358 414 SUSHI 6.
FT DISULFID 58 103 BY SIMILARITY.
FT DISULFID 88 115 BY SIMILARITY.
FT DISULFID 120 160 BY SIMILARITY.
FT DISULFID 146 176 BY SIMILARITY.
FT DISULFID 181 223 BY SIMILARITY.
FT DISULFID 209 240 BY SIMILARITY.
FT DISULFID 245 287 BY SIMILARITY.
FT DISULFID 273 299 BY SIMILARITY.
FT DISULFID 303 343 BY SIMILARITY.
FT DISULFID 329 355 BY SIMILARITY.
FT DISULFID 359 400 BY SIMILARITY.
FT DISULFID 386 413 BY SIMILARITY.
FT CARBOHYD 74 74 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 275 275 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 366 366 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 428 428 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 469 AA: 51551 MW: 41E137CB8D8C6321 CRC64:

Query Match 19.7%; Score 144.5; DB 1; Length 469;
Best Local Similarity 29.4%; Pred. No. 6.1e-07;
Matches 32; Conservative 19; Mismatches 49; Indels 9; Gaps 3;

OY 22 SFSSQCAFSCSEGTNLGIEETTCGPPGN---WSPEPTCQVIOCEPLAPDGIIMNC 78
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 137 TFGSQIEFSCQEGHILVGSSTSCSEVKGAVAMNSPPECVIYKGP--PPDISNKHSG 194
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
OY 79 PLASFSTACFTICSGTELGKKTIC---ESSGIWNSPPIQOKL 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 195 TEDFPYNHGISTYCDPGRFVGSFPGICTVNKTVPVWSSPPTCEKI 243
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 33
C4BP_RAT STANDARD: PRT: 258 AA.
ID C4BP_RAT
AC Q63515:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE C4b-binding protein beta chain precursor.
GN C4BPB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;
RX MEDLINE=9716082; PubMed=9013975;
RA Hillarp A., Wiklund H., Thern A., Dahlback B.;
RT "Molecular cloning of rat C4b binding protein alpha- and beta-chains:
  structural and functional relationships among human, bovine, rabbit,
  mouse, and rat proteins."
RL J. Immunol. 158:1315-1323(1997).
CC -1- FUNCTION: C4BP CONTROLS THE CLASSICAL PATHWAY OF COMPLEMENT
  ACTIVATION. IT BINDS AS A COFACTOR TO C3b/C4b INACTIVATOR
  (C3bINA), WHICH THEN HYDROLYZES THE COMPLEMENT FRAGMENT C4b. IT
  ALSO ACCELERATES THE DEGRADATION OF THE C4bC2a COMPLEX (C3
  CONVERTASE) BY DISSOCIATING THE COMPLEMENT FRAGMENT C2a. IT
  INTERACTS ALSO WITH ANTICOAGULANT PROTEIN S AND WITH SERUM AMYLOID
  P COMPONENT.
CC -1- SUBUNIT: DISULFIDE-LINKED COMPLEX OF C4BP ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: TO C4BP ALPHA CHAIN AND TO PIG APOLIPOPROTEIN R.
CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

```

```

CC      DISSOCIATION OF C3 CONVERTASE.
CC      -1- SUBCELLULAR LOCATION: CLASS GPI: ATTACHED TO THE MEMBRANE BY A
CC          GPI-ANCHOR.
CC      -1- ALTERNATIVE PRODUCTS: THERE ARE AT LEAST SIX FORMS OF DAF
CC          TRANSCRIPTS (DISTRIBUTED IN FOUR CLASSES: GPI, TCL, TCS AND SEC)
CC          GENERATED BY ALTERNATIVE SPLICING. ONE GPI-ANCHORED FORM (GDAB-
CC          GPI), TWO FORMS WITH LONGER CYTOPLASMIC REGION (GDAB-TCL AND
CC          GDAB-C-TCL), TWO FORMS WITH SHORTER CYTOPLASMIC REGION (GDA-TCS AND
CC          GDAB-TCS) AND ONE SECRETED FORM (GDAB-SEC).
CC      -1- TISSUE SPECIFICITY: ALL THE ISOFORMS ARE WIDELY EXPRESSED. GPI AND
CC          TCS ARE THE MAJOR FORMS, WHEREAS SEC IS MINOR AND TCL IS ONLY
CC          PRESENT IN TRACE LEVELS.
CC      -1- SIMILARITY: CONTAINS 4 SUSHI (SCF) DOMAINS.
CC      -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
CC          (RCA) FAMILY.
CC      -----
CC      THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; D55667; BAA09514.1; JOINED.
DR      EMBL; D55656; BAA09514.1; JOINED.
DR      EMBL; D55657; BAA09514.1; JOINED.
DR      EMBL; D55658; BAA09514.1; JOINED.
DR      EMBL; D55659; BAA09514.1; JOINED.
DR      EMBL; D55660; BAA09514.1; JOINED.
DR      EMBL; D55661; BAA09514.1; JOINED.

```

DR	EMBL	D55664	BAA09514.1	JOINED
DR	EMBL	D55667	BAA09515.1	-
DR	EMBL	D55666	BAA09515.1	JOINED
DR	EMBL	D55657	BAA09515.1	JOINED
DR	EMBL	D55658	BAA09515.1	JOINED
DR	EMBL	D55659	BAA09515.1	JOINED
DR	EMBL	D55660	BAA09515.1	JOINED
DR	EMBL	D55661	BAA09515.1	JOINED
DR	EMBL	D55662	BAA09515.1	JOINED
DR	EMBL	D55663	BAA09515.1	JOINED
DR	EMBL	D55664	BAA09515.1	JOINED
DR	EMBL	D55665	BAA09515.1	JOINED
DR	EMBL	D55667	BAA09516.1	-
DR	EMBL	D55656	BAA09516.1	JOINED
DR	EMBL	D55657	BAA09516.1	JOINED
DR	EMBL	D55658	BAA09516.1	JOINED
DR	EMBL	D55659	BAA09516.1	JOINED
DR	EMBL	D55660	BAA09516.1	JOINED
DR	EMBL	D55661	BAA09516.1	JOINED
DR	EMBL	D55662	BAA09516.1	JOINED
DR	EMBL	D55663	BAA09516.1	JOINED
DR	EMBL	D55664	BAA09516.1	JOINED
DR	EMBL	D55665	BAA09516.1	JOINED
DR	EMBL	D55667	BAA09517.1	-
DR	EMBL	D55656	BAA09517.1	JOINED
DR	EMBL	D55657	BAA09517.1	JOINED
DR	EMBL	D55658	BAA09517.1	JOINED
DR	EMBL	D55659	BAA09517.1	JOINED
DR	EMBL	D55660	BAA09517.1	JOINED
DR	EMBL	D55661	BAA09517.1	JOINED
DR	EMBL	D55662	BAA09517.1	JOINED
DR	EMBL	D55663	BAA09517.1	JOINED
DR	EMBL	D55664	BAA09517.1	JOINED
DR	EMBL	D55665	BAA09517.1	JOINED
DR	EMBL	D55667	BAA09518.1	-
DR	EMBL	D55656	BAA09518.1	JOINED
DR	EMBL	D55657	BAA09518.1	JOINED
DR	EMBL	D55658	BAA09518.1	JOINED
DR	EMBL	D55659	BAA09518.1	JOINED





DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00084; sush1. 4.  
 DR SMART: SM00032; CCP; 4.  
 KW Heparin-binding; Glycoprotein; Plasma; Repeat; Sush1; Signal.  
 FT SIGNAL 1 19  
 FT CHAIN 20 345  
 FT DOMAIN 22 80 BETA-2-GLYCOPROTEIN I.  
 FT DOMAIN 83 138 SUSHI 1.  
 FT DOMAIN 141 201 SUSHI 2.  
 FT DOMAIN 204 261 SUSHI 3.  
 FT DOMAIN 263 345 SUSHI 4.  
 FT DOMAIN 263 345 SUSHI-LIKE.  
 FT DISULFID 23 66  
 FT DISULFID 51 79  
 FT DISULFID 84 124  
 FT DISULFID 110 137  
 FT DISULFID 142 188  
 FT DISULFID 174 200  
 FT DISULFID 205 248  
 FT DISULFID 234 260  
 FT DISULFID 264 315  
 FT DISULFID 300 325  
 FT DISULFID 307 345  
 FT CARBOHYD 92 92  
 FT CARBOHYD 162 162  
 FT CARBOHYD 183 183  
 FT CARBOHYD 193 193  
 FT CARBOHYD 253 253  
 FT CONFLICT 101 101  
 FT CONFLICT 108 108  
 FT CONFLICT 177 177  
 FT CONFLICT 191 191  
 FT CONFLICT 194 194  
 FT CONFLICT 259 259  
 FT CONFLICT 302 302  
 FT CONFLICT 305 305  
 FT CONFLICT 329 329  
 SQ SEQUENCE 345 AA; 38252 MW; E117DAB609461C33 CRC64;  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 N-LINKED (GLCNAC. . .)  
 E -> G (IN REF. 1).  
 F -> S (IN REF. 1).  
 H -> R (IN REF. 1).  
 H -> C (IN REF. 3).  
 W -> C (IN REF. 1).  
 S -> N (IN REF. 1).  
 H -> N (IN REF. 1).  
 K -> R (IN REF. 1).  
 H -> R (IN REF. 1).

Query Match 19.3%; Score 141.5; DB 1; Length 345;  
 Best Local Similarity 31.5%; Pred. No. 8.7e-07;  
 Matches 39; Conservative 15; Mismatches 63; Indels 7; Gaps 6;  
 OY 1 CEPLAEPELTMDCTHPF-GNFSF-SSQCAFSCSEGTNLGIEETTCGPFGNMSPEPIC 58  
 DB 142 CEPPPIPKRASLSYVKFLGNSFYSKAVFKCLPHHAFMGNDTVCTEHNMTQ-LPEC 200  
 OY 59 QVIOCEPLSAPDLGIMNCSHPL-ASFSTSACTFICSEGTTELIGKKTKICSSGIVSNPS 117  
 DB 201 REVRCFPSPRPDNGFVN-HRANPVLYKDTATFGCHETYSLDGPREVECSKFGNMS-AQ 257  
 OY 118 PTIQ 121  
 DB 258 PSCK 261

RESULT 36  
 CR2\_HUMAN  
 ID CR2\_HUMAN STANDARD; PRT; 1033 AA.  
 AC P20023;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Complement receptor type 2 precursor (Cr2) (Complement C3d receptor)  
 DE (Epstein-Barr virus receptor) (EBV receptor) (CD21 antigen).  
 GN CR2 OR C3DR.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Euteria; Primates; Catarrhini; Homindae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89123377; PubMed=2563370;

RA Fujisaku A., Harley J.B., Frank M.B., Gruner B.A., Frazier B.,  
 RA Holers V.M.;  
 RT "Genomic organization and polymorphisms of the human C3d/Epstein-Barr  
 RT virus receptor";  
 RL J. Biol. Chem. 264:2118-2125(1989).  
 RN [2]  
 RP SEQUENCE OF 226-233; 256-267; 332-341; 667-677 AND 898-908.  
 RX MEDLINE=86287311; PubMed=3016712;  
 RA Wels J.J., Pearson D.T., Klickstein L.B., Wong W.W., Richards S.A.,  
 RA de Bruyn Kops A., Smith J.A., Wels J.H.;  
 RT "Identification of a partial cDNA clone for the C3d/Epstein-Barr  
 RT virus receptor of human B lymphocytes: homology with the receptor for  
 RT fragments C3b and C4b of the third and fourth components of  
 RT complement";  
 RL Proc. Natl. Acad. Sci. U.S.A. 83:5639-5643(1986).  
 RN [3]  
 RP SEQUENCE OF 492-556 FROM N.A. (SHORT ISOFORM).  
 RX MEDLINE=93294286; PubMed=8390533;  
 RA Sinha S.K., Todd S.C., Hedrick J.A., Spelsberg C.L., Lambiris J.D.,  
 RA Tsoukas C.D.;  
 RT "Characterization of the EBV/C3d receptor on the human Jurkat T cell  
 RT line: evidence for a novel transcript";  
 RL J. Immunol. 150:5311-5320(1993).  
 CC - FUNCTION: RECEPTOR FOR COMPLEMENT C3D AND FOR THE EPSTEIN-BARR  
 CC VIRUS ON HUMAN B-CELLS AND T-CELLS. PARTICIPATES IN B LYMPHOCYTES  
 CC ACTIVATION.  
 CC - SUBCELLULAR LOCATION: Type I membrane protein.  
 CC - ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC - TISSUE SPECIFICITY: MATURE B LYMPHOCYTES, T LYMPHOCYTES AND  
 CC FOLLICULAR DENDRITIC CELLS OF THE SPLEEN.  
 CC - SIMILARITY: TO MOUSE CR2. CD21 IS A MEMBER OF THE REGULATORS OF  
 CC COMPLEMENT ACTIVATION (RCA) FAMILY.  
 CC - SIMILARITY: CONTAINS 15 SUSHI (SCR) DOMAINS.  
 CC - DATABASE: NAME=PROW; NOTE=CD guide CD21 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd21.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib.ch](mailto:license@sib.ch)).  
 CC -----  
 DR EMBL: M26004; AAA35786.1; -;  
 DR EMBL: M26016; AAB04638.1; -;  
 DR EMBL: M24007; AAB04638.1; JOINED.  
 DR EMBL: M24008; AAB04638.1; JOINED.  
 DR EMBL: M24009; AAB04638.1; JOINED.  
 DR EMBL: M24010; AAB04638.1; JOINED.  
 DR EMBL: M24011; AAB04638.1; JOINED.  
 DR EMBL: M26009; AAB04638.1; JOINED.  
 DR EMBL: M26010; AAB04638.1; JOINED.  
 DR EMBL: M26011; AAB04638.1; JOINED.  
 DR EMBL: M26012; AAB04638.1; JOINED.  
 DR EMBL: M26013; AAB04638.1; JOINED.  
 DR EMBL: M26014; AAB04638.1; JOINED.  
 DR EMBL: M26015; AAB04638.1; JOINED.  
 DR EMBL: S62696; AAB27186.1; -;  
 DR PIR: A32036; A32036.  
 DR PIR: A24319; A24319.  
 DR PIR: B24319; B24319.  
 DR PIR: C24319; C24319.  
 DR PIR: D24319; D24319.  
 DR PIR: E24319; E24319.  
 DR PIR: F24319; F24319.  
 DR HSSP: P10998; 1YVD.  
 DR MIM: 120650; -;  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF00084; sush1. 15.  
 DR SMART: SM00032; CCP; 14.  
 KW Complement pathway; Glycoprotein; Transmembrane; Repeat; Signal;

FT	KM	Receptor; Sushi; Alternative splicing.	1	20	
FT	FT	SIGNAL	21	1033	COMPLEMENT RECEPTOR TYPE 2.
FT	FT	CHAIN	21	971	EXTRACELLULAR (POTENTIAL).
FT	FT	DOMAIN	972	999	POTENTIAL.
FT	FT	TRANSMEM	1000	1033	CYTOSOLASMIC (POTENTIAL).
FT	FT	DOMAIN	22	83	SUSHI 1.
FT	FT	DOMAIN	90	147	SUSHI 2.
FT	FT	DOMAIN	153	211	SUSHI 3.
FT	FT	DOMAIN	214	272	SUSHI 4.
FT	FT	DOMAIN	275	343	SUSHI 5.
FT	FT	DOMAIN	350	407	SUSHI 6.
FT	FT	DOMAIN	409	467	SUSHI 7.
FT	FT	DOMAIN	470	523	SUSHI 8.
FT	FT	DOMAIN	526	594	SUSHI 9.
FT	FT	DOMAIN	601	658	SUSHI 10.
FT	FT	DOMAIN	661	715	SUSHI 11.
FT	FT	DOMAIN	718	780	SUSHI 12.
FT	FT	DOMAIN	787	844	SUSHI 13.
FT	FT	DOMAIN	850	908	SUSHI 14.
FT	FT	DOMAIN	911	969	SUSHI 15.
FT	FT	DOMAIN	23	65	BY SIMILARITY.
FT	FT	DISULFID	51	82	BY SIMILARITY.
FT	FT	DISULFID	91	132	BY SIMILARITY.
FT	FT	DISULFID	118	146	BY SIMILARITY.
FT	FT	DISULFID	154	197	BY SIMILARITY.
FT	FT	DISULFID	183	210	BY SIMILARITY.
FT	FT	DISULFID	215	256	BY SIMILARITY.
FT	FT	DISULFID	242	271	BY SIMILARITY.
FT	FT	DISULFID	276	325	BY SIMILARITY.
FT	FT	DISULFID	305	342	BY SIMILARITY.
FT	FT	DISULFID	351	393	BY SIMILARITY.
FT	FT	DISULFID	379	406	BY SIMILARITY.
FT	FT	DISULFID	410	453	BY SIMILARITY.
FT	FT	DISULFID	430	466	BY SIMILARITY.
FT	FT	DISULFID	471	509	BY SIMILARITY.
FT	FT	DISULFID	495	522	BY SIMILARITY.
FT	FT	DISULFID	527	576	BY SIMILARITY.
FT	FT	DISULFID	556	593	BY SIMILARITY.
FT	FT	DISULFID	602	644	BY SIMILARITY.
FT	FT	DISULFID	630	657	BY SIMILARITY.
FT	FT	DISULFID	662	699	BY SIMILARITY.
FT	FT	DISULFID	685	714	BY SIMILARITY.
FT	FT	DISULFID	719	762	BY SIMILARITY.
FT	FT	DISULFID	748	779	BY SIMILARITY.
FT	FT	DISULFID	788	830	BY SIMILARITY.
FT	FT	DISULFID	816	843	BY SIMILARITY.
FT	FT	DISULFID	851	894	BY SIMILARITY.
FT	FT	DISULFID	880	907	BY SIMILARITY.
FT	FT	DISULFID	912	955	BY SIMILARITY.
FT	FT	DISULFID	941	968	BY SIMILARITY.
FT	FT	CARBOHYD	121	121	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	127	127	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	294	294	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	372	372	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	492	492	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	623	623	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	682	682	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	800	800	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	823	823	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	861	861	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	CARBOHYD	911	911	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	FT	VANSPPLIC	499	524	MISSING (IN SHORT ISOFORM).
FT	FT	VANSPPLIC	525	556	TCGCPPVYVNAHGHSLEDPVCTVTVTC -> NHLPPI
FT	FT	VANSPPLIC			TCQYLDWMGTHRFELFRFSIMNHGLHM (IN SHORT
FT	FT	VANSPPLIC			ISOFORM).
FT	FT	CONFLICT	667	667	Q -> D (IN REF. 2).
FT	FT	CONFLICT	902	902	Q -> G (IN REF. 2).
FT	FT	CONFLICT	906	906	H -> L (IN REF. 2).
SQ	SEQUENCE	1033 AA; 112973 MM; 1749DBA07847ADA CRC64;			
Query Match					
19.28; Score 141; DB 1; Length 1033;					

Best Local Similarity 29.4%; Pred. No. 3e-06;  
Matches 35; Conservative 16; Mismatches 62; Indels 6; Gaps

7 PELGTMDCTHFGCN---SFSSQCAFCSCSPGNTWTCIEETTGPGPNMSSPEPTCYIOC 64  
Db 853 PPKPTKPNNGHHGNIAFNFSFGMSALIVSCQGYLVNGEALLTCHGTWSPAPRCKRYNC 91  
Qy 64 EPLSAPDL-GIMNCSHPLASFSTTSACTPTICSEGTLEIGKKTKTICSSGIVNSPDIQ 121  
Db 913 S--SEADMDGIQKGLERPKMYGYGAVVLTCEDEGTYLEGSPOSOSPDHQMNPPLAVCR 969  
RESULT 37  
CFAB\_HUMAN STANDARD; PRT: 764 AA.  
ID P00751; O15006; Q29944; Q9BTF5; Q9BX92;  
AC 21-JUL-1986 (Rel. 01, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Complement factor B precursor (EC 3.4.21.47) (C3/C5 convertase)  
DE (Properdin factor B) (Glycine-rich beta glycoprotein) (GBG) (PBF2).  
GN BF.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxId=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ALLELES S; FA AND FB).  
RX MEDLINE=91065702; PubMed=2249879;  
RA Davrinche C., Abbel M., Clerc A.;  
RT "Molecular characterization of human complement factor B subtypes";  
RL Immunogenetics 32:309-312(1990).  
RN [2]  
RP SEQUENCE FROM N.A. (ALLELE S).  
RC TISSUE=Liver;  
RX MEDLINE=94237735; PubMed=8181962;  
RA Mejia J.E., Jahn I., de la Salla H., Hauptmann G.;  
RT "Human factor B. Complete cDNA sequence of the BF\*s allele.";  
RL Hum. Immunol. 39:49-53(1994).  
RN [3]  
RP SEQUENCE FROM N.A. (ALLELE S).  
RC TISSUE=Liver;  
RX MEDLINE=94041399; PubMed=8225386;  
RA Schwabe W., Luettig B., Sokolowski T., Estaller C., Weiss E.H.,  
RT Meyer Zum Bueschenfelde K.-H., Walely K., Dipplod W.;  
RT "Human complement factor B: functional properties of a recombinant  
RT zymogen of the alternative activation pathway convertase.";  
RL Immunobiology 188:221-232(1993).  
RN [4]  
RP SEQUENCE FROM N.A. (ALLELE S).  
RX MEDLINE=94067177; PubMed=8247029;  
RA Horiuchi T., Kim S., Matsumoto M., Watanabe I., Fujita S.,  
RT Volanakis J.E.;  
RT "Human complement factor B: cDNA cloning, nucleotide sequencing,  
RT phenotypic conversion by site-directed mutagenesis and expression.";  
RL Mol. Immunol. 30:1587-1592(1993).  
RN [5]  
RP SEQUENCE FROM N.A.  
RX Rowen L., Dankers C., Baskin D., Faust J., Loretz C., Ahearn M.E.,  
RA Bata A., Swartzell S., Smith T.M., Spies T., Hood L.;  
RT "Sequence determination of 300 kilobases of the human class III MHC  
RT locus.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Colon;  
RA Strausberg R.;  
RT Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
RN [7]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RA Jaatien T., Kanerva J., Poutanen K.E., Saarinen-Pihkala U.,  
RT Lokki M.-L.;  
RT "Expression and alternative splicing of human factor B gene in

RT leukemic mononuclear cells.";  
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.  
RN [8]  
RP SEQUENCE OF 26-764, PARTIAL SEQUENCE FROM N.A., AND CARBOHYDRATES.  
RX MEDLINE-84161997; PubMed-6546754;  
RA Mole J.E., Anderson J.K., Davison E.A., Woods D.E.;  
RT "Complete primary structure for the zymogen of human complement  
RT factor B.";  
RL J. Biol. Chem. 259:3407-3412(1984).  
RN [9]  
RP SEQUENCE OF 260-764.  
RX MEDLINE-83204002; PubMed-6342610;  
RA Christie D.L., Gagnon J.;  
RT "Amino acid sequence of the Bb fragment from complement factor B.  
RT Sequence of the major cyanogen bromide-cleavage peptide (CB-II) and  
RL completion of the sequence of the Bb fragment.";  
RL Biochem. J. 209:61-70(1983).  
RN [10]  
RP SEQUENCE OF 339-764 FROM N.A.  
RX MEDLINE-83273641; PubMed-6308626;  
RA Campbell R.D., Porter R.R.;  
RT "Molecular cloning and characterization of the gene coding for human  
RT complement protein factor B.";  
RL Proc. Natl. Acad. Sci. U.S.A. 80:4464-4468(1983).  
RN [11]  
RP SEQUENCE OF 467-595 AND 752-764 FROM N.A.  
RX MEDLINE-83039428; PubMed-6957884;  
RA Woods D.E., Markham A.F., Ricker A.T., Goldberger G., Colten H.R.;  
RT "Isolation of cDNA clones for the human complement protein factor B,  
RT a class III major histocompatibility complex gene product.";  
RL Proc. Natl. Acad. Sci. U.S.A. 79:5661-5665(1982).  
RN [12]  
RP SEQUENCE OF 16-259 FROM N.A.  
RX MEDLINE-84158524; PubMed-6523161;  
RA Morley B.J., Campbell R.D.;  
RT "Internal homologues of the Ba fragment from human complement  
RT component factor B, a class III MHC antigen.";  
RL EMBO J. 3:153-157(1984).  
RN [13]  
RP SEQUENCE OF 1-99 FROM N.A.  
RX TISSUE-Blood;  
RC MEDLINE-87102880; PubMed-3643061;  
RA Wu L.C., Morley B.J., Campbell R.D.;  
RT "Cell-specific expression of the human complement protein factor B  
RT gene: evidence for the role of two distinct 5'-flanking elements.";  
RL Cell 48:331-342(1987).  
RN [14]  
RP GLYCATION IN POSITION 291.  
RX MEDLINE-91174758; PubMed-2006911;  
RA Niemann M.A., Bhown A.S., Miller E.J.;  
RT "The principal site of glycation of human complement factor B.";  
RL Biochem. J. 274:473-480(1991).  
CC -1- FUNCTION: FACTOR B WHICH IS PART OF THE ALTERNATE PATHWAY OF THE  
CC COMPLEMENT SYSTEM IS CLEAVED BY FACTOR D INTO 2 FRAGMENTS: BA AND  
CC BB. BA SERINE PROTEASE, THEN COMBINES WITH COMPLEMENT FACTOR 3B  
CC TO GENERATE THE C3 OR C5 CONVERTASE. IT HAS ALSO BEEN IMPLICATED  
CC IN PROLIFERATION AND DIFFERENTIATION OF PREACTIVATED B  
CC LYMPHOCYTES, RAPID SPREADING OF PERIPHERAL BLOOD MONOCYTES,  
CC STIMULATION OF LYMPHOCYTE BLASTOGENESIS AND LYSIS OF ERYTHROCYTES.  
CC BA INHIBITS THE PROLIFERATION OF PREACTIVATED B LYMPHOCYTES.  
CC -1- CATALYTIC ACTIVITY: Cleaves C3 in the alpha-chain to yield C3a and  
CC C3b. Cleaves C5 in the alpha-chain to yield C5a and C5b. Both  
CC cleavages take place at the C-terminal of an arginine residue.  
CC -1- SUBUNIT: MONOMER.  
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are  
CC produced by alternative splicing.  
CC -1- POLYMORPHISM: TWO MAJOR VARIANTS, F AND S, AND 2 MINOR VARIANTS,  
CC AS WELL AS AT LEAST 14 VERY RARE VARIANTS, HAVE BEEN IDENTIFIED.  
CC -1- MISCELLANEOUS: FACTOR B IS A MAJOR HISTOCOMPATIBILITY COMPLEX  
CC CLASS-III PROTEIN.  
CC -1- SIMILARITY: WITH COMPLEMENT C2.  
CC -1- SIMILARITY: CONTAINS 3 SUSHI (SCR) DOMAINS.  
CC -1- SIMILARITY: CONTAINS 1 VWFA DOMAIN.

CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S1; ALSO KNOWN AS THE  
CC TRYPSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: X72875; CAA51389.1; -  
CC EMBL: S67310; AAD13989.1; -  
CC EMBL: L15702; AAA16820.1; -  
CC EMBL: X00284; CAA25077.1; -  
CC EMBL: AF019413; AAB67977.1; -  
CC EMBL: BC004143; AAH04143.1; -  
CC EMBL: AF349679; AAK30167.1; -  
CC EMBL: K01566; AAA36225.2; -  
CC EMBL: J00125; -; NOT\_ANNOTATED\_CDS.  
CC EMBL: J00126; AAA36276.1; -  
CC EMBL: J00185; AAA36219.1; ALT\_SEQ.  
CC EMBL: J00186; AAA36220.1; -  
CC EMBL: M15082; AAA59625.1; -  
CC PIR: A00934; BBHU.  
CC PIR: S14339; S14339.  
CC PIR: S34075; S34075.  
CC HSP: P00734; 2HNT.  
CC MEROPS: S01.196; -  
CC SWISS-2DPAGE: P00751; HUMAN.  
CC S1ema-2DPAGE: P00751; -  
CC M138470; -  
CC InterPro: IPR001314; Chymotrypsin.  
CC InterPro: IPR000436; Susli\_SCR\_CCP.  
CC InterPro: IPR001254; Trypsin.  
CC InterPro: IPR002035; VWFA.  
CC Pfam: PF00084; susli; 3.  
CC Pfam: PF00089; trypsin; 1.  
CC Pfam: PF00092; vwa; 1.  
CC PRINTS: PR00722; CHYMOTRYPSIN.  
CC PRINTS: PR00453; VWFADOMAIN.  
CC SMART: SM00032; CCP; 3.  
CC SMART: SM00020; TRYP\_SPE; 1.  
CC SMART: SM00327; VWA; 1.  
CC PROSITE: PS50240; TRYPSIN\_DOM; 1.  
CC PROSITE: PS00134; TRYPSIN\_HTS; 1.  
CC PROSITE: PS00135; TRYPSIN\_SER; 1.  
CC PROSITE: PS50234; VWFA; 1.  
CC Complement alternate pathway; Plasma; Hydrolase; Serine protease;  
CC Glycoprotein; Repeat; Susli; Signal; Polymorphism; Zymogen;  
CC Alternative splicing.  
CC FT SIGNAL 1 25  
CC FT CHAIN 26 764  
CC FT CHAIN 26 259  
CC FT CHAIN 260 764  
CC FT DOMAIN 36 99  
CC FT DOMAIN 102 159  
CC FT DOMAIN 164 219  
CC FT DOMAIN 270 469  
CC FT DOMAIN 482 764  
CC FT ACT\_SITE 526 526  
CC FT ACT\_SITE 576 576  
CC FT ACT\_SITE 699 699  
CC FT ACT\_SITE 37 76  
CC FT DISULFID 62 98  
CC FT DISULFID 103 145  
CC FT DISULFID 131 158  
CC FT DISULFID 165 205  
CC  
CC Query Match 19.2%; Score 140.5; DB 1; Length 764;  
CC Best local Similarity 25.2%; Pred. No. 2.5e-06;  
CC Matches 35; Conservative 14; Mismatches 55; Indels 35; Gaps 4;



RX MEDLINE=87115845; PubMed=2433596;  
 RA Caras I.W., Davitz M.A., Rhee L., Meddell G., Martin D.W. Jr.,  
 RA Nussenzweig V.;  
 RT "Cloning of decay-accelerating factor suggests novel use of splicing  
 RT to generate two proteins.";  
 RL Nature 325:545-549(1987).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Cervix;  
 RA Strausberg R.;  
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RN SEQUENCE OF 6-381 FROM N.A. (ISOFORM 2).  
 RX MEDLINE=87175602; PubMed=2436222;  
 RA Medof M.E., Lublin D.M., Holers V.M., Ayers D.J., Getty R.R.,  
 RA Leykam J.F., Atkinson J.P., Tykocinski M.L.;  
 RT "Cloning and characterization of cDNAs encoding the complete sequence  
 RT of decay-accelerating factor of human complement.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 84:2007-2011(1987).  
 RN [4]  
 RN SEQUENCE OF 35-381 FROM N.A. (ISOFORM 2).  
 RC TISSUE=Hippocampus;  
 RA Kumar V.B., Hyung C., Nakra R., Walters M., Sasser T., Bernardo A.;  
 RT "Decay-acceleration factor (DAF; CD 55) in the brain of Alzheimer's  
 RT disease patients.";  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RN SEQUENCE OF 1-100 FROM N.A.  
 RX MEDLINE=91271256; PubMed=1711208;  
 RA Ewulonu U.K., Ravi L., Medof M.E.;  
 RT "Characterization of the decay-accelerating factor gene promoter  
 RT region.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:4675-4679(1991).  
 RN [6]  
 RN SEQUENCE OF 35-46.  
 RC TISSUE=Urine;  
 RX MEDLINE=91291869; PubMed=1712233;  
 RA Nakano Y., Sugita Y., Ishikawa Y., Choi N.-H., Tobe T., Tomita M.;  
 RT "Isolation of two forms of decay-accelerating factor (DAF) from human  
 RT urine.";  
 RL Biochim. Biophys. Acta 1074:326-330(1991).  
 RN [7]  
 RN GPI-ANCHOR.  
 RX MEDLINE=91093238; PubMed=1824699;  
 RA Moran P., Raab H., Kohr W.J., Caras I.W.;  
 RT "Glycophospholipid membrane anchor attachment. Molecular analysis of  
 RT the cleavage/attachment site.";  
 RL J. Biol. Chem. 266:1250-1257(1991).  
 RN [8]  
 RN DISULFIDE BONDS IN SUSHI DOMAINS.  
 RX MEDLINE=92305034; PubMed=1377029;  
 RA Nakano Y., Sumida K., Kikuta N., Miura N.-H., Tobe T., Tomita M.;  
 RT "Complete determination of disulfide bonds localized within the short  
 RT consensus repeat units of decay accelerating factor (CD55 antigen).";  
 RL Biochim. Biophys. Acta 1116:235-240(1992).  
 RN [9]  
 RN FUNCTION AS A ECHOVIRUS RECEPTOR.  
 RX MEDLINE=95045399; PubMed=7525274;  
 RA Ward T., Pipkin P.A., Clarkson N.A., Stone D.M., Minor P.D.,  
 RA Almond J.W.;  
 RT "Decay-accelerating factor CD55 is identified as the receptor for  
 RT echovirus 7 using CELICs, a rapid immuno-focal cloning method.";  
 RL EMO J. 13:5070-5074(1994).  
 RN [10]  
 RN VARIANT BLOOD GROUP DR(A-).  
 RX MEDLINE=94325573; PubMed=7519480;  
 RA Lublin D.M., Mallinson G., Poole J., Reid M.E., Thompson E.S.,  
 RA Ferdman B.R., Telen M.J., Anstee D.J., Tanner M.J.A.;  
 RT "Molecular basis of reduced or absent expression of  
 RT decay-accelerating factor in Cromer blood group phenotypes.";  
 RL Blood 84:1276-1282(1994).  
 CC -1- FUNCTION: THIS PROTEIN RECOGNIZES C4B AND C3B FRAGMENTS THAT  
 CC CONDENSE WITH CELL-SURFACE HYDROXYL OR AMINO GROUPS WHEN NASCENT

CC C4B AND C3B ARE LOCALLY GENERATED DURING C4 AND C3 ACTIVATION.  
 CC INTERACTION OF DAF WITH CELL-ASSOCIATED C4B AND C3B POLYPEPTIDES  
 CC INTERFERES WITH THEIR ABILITY TO CATALYZE THE CONVERSION OF C2 AND  
 CC FACTOR B TO ENZYMATICALLY ACTIVE C2A AND Bb AND THEREBY PREVENTS  
 CC THE FORMATION OF C4B2A AND C3BB, THE AMPLIFICATION CONVERTASES OF  
 CC THE COMPLEMENT CASCADE.  
 CC -1- FUNCTION: ALSO ACTS AS THE RECEPTOR FOR ECHOVIRUS 7 AND RELATED  
 CC VIRUSES (ECHOVIRUSES 13, 21, 29 AND 33).  
 CC -1- SUBUNIT: MONOMER (MAJOR FORM) AND NON-DISULFIDE-LINKED, COVALENT  
 CC HOMODIMER (MINOR FORM).  
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS, 1/DAF-1 AND 2/DAF-2 (SHOWN  
 CC HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -1- TISSUE SPECIFICITY: EXPRESSED ON THE PLASMA MEMBRANES OF ALL CELL  
 CC TYPES THAT ARE IN INTIMATE CONTACT WITH PLASMA COMPLEMENT  
 CC PROTEINS. IT IS ALSO FOUND ON THE SURFACES OF EPITHELIAL CELLS  
 CC Lining EXTRACELLULAR COMPARTMENTS, AND VARIANTS OF THE MOLECULE  
 CC ARE PRESENT IN BODY FLUIDS AND IN EXTRACELLULAR MATRIX.  
 CC -1- DOMAIN: THE FIRST SUSHI DOMAIN (SCR1) IS NOT NECESSARY FOR  
 CC FUNCTION. SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE  
 CC ACTIVE SITE ON SCR3 (BY SIMILARITY).  
 CC -1- PTR: THE SER/THR-RICH DOMAIN IS HEAVILY O-GLYCOSYLATED.  
 CC -1- POLYMORPHISM: DAF IS RESPONSIBLE FOR THE CROMER BLOOD GROUP  
 CC SYSTEM. IT CONSISTS OF AT LEAST SEVEN HIGH-INCIDENCE (CR(A),  
 CC TC(A), DR(A), ES(A), WES(B), UMC, AND IFC) AND LOW-INCIDENCE  
 CC (TC(B), TC(C), AND WES(A)) ANTIGENS THAT RESIDE ON DAF. IN THE  
 CC CROMER PHENOTYPES DR(A-) AND INAB THERE IS REDUCED OR ABSENT  
 CC EXPRESSION OF DAF, RESPECTIVELY. IN THE CASE OF THE DR(A-) PHENOTYPE,  
 CC A SINGLE NUCLEOTIDE SUBSTITUTION WITHIN EXON 5 ACCOUNTS  
 CC FOR TWO CHANGES: A SIMPLE AMINO ACID SUBSTITUTION THAT IS THE  
 CC BASIS OF THE ANTIGENIC VARIATION, AND AN ALTERNATIVE SPLICING  
 CC EVENT THAT UNDERLIES THE DECREASED EXPRESSION OF DAF IN THIS  
 CC PHENOTYPE.  
 CC -1- SIMILARITY: CONTRAINS 4 SUSHI (SCR) DOMAINS.  
 CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION  
 CC (RCA) FAMILY.  
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD55 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prow/cd/cd55.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL, M31516; AAA52169.1; -;  
 CC DR EMBL, M30142; AAA52168.1; -;  
 CC DR EMBL, BC001288; AAH01288.1; -;  
 CC DR EMBL, M15799; AAA52167.1; -;  
 CC DR EMBL, U88576; AAB48692.1; -;  
 CC DR EMBL, M64553; AAA52170.1; -;  
 CC DR EMBL, M64366; AAA52170.1; JOINED.  
 CC DR EMBL, S72858; AAC60633.1; -;  
 CC DR PIR, B26359; B26359.  
 CC DR PIR, A26359; A26359.  
 CC DR PIR, S16187; S16187.  
 CC DR PIR, A39101; A39101.  
 CC DR PIR, S23138; S23138.  
 CC DR HSSP, P08603; 1HCC.  
 CC DR MIM, 125240; -;  
 CC DR InterPro: IPR000436; Sushi\_SCR\_CCP.  
 CC DR Pfam: PF00084; sushi\_4.  
 CC DR SMART: SM00032; CCP; 4.  
 CC KW Complement pathway; Plasma; glycoprotein; Membrane; Repeat;  
 CC KW Alternative splicing; GPI-anchor; Signal; Sushi; Polymorphism;  
 CC KW Blood group antigen.  
 CC FT SIGNAL 1 34  
 CC FT CHAIN 35 353 COMPLEMENT DECAY-ACCELERATING FACTOR.  
 CC FT PROPEP 354 381 REMOVED IN MATURE FORM.  
 CC FT DOMAIN 35 95 SUSHI 1.  
 CC FT 97 159 SUSHI 2.

```

FT DOMAIN 162 221 SUSHI 3.
FT DOMAIN 224 284 SUSHI 4.
FT DOMAIN 287 356 SER/THR-RICH.
FT DISULFID 36 81
FT DISULFID 65 94
FT DISULFID 98 145
FT DISULFID 129 158
FT DISULFID 163 204
FT DISULFID 190 220
FT DISULFID 225 267
FT DISULFID 253 283
FT CARBOHYD 95 95
FT CARBOHYD 353 353
FT LIPID 362 381
FT VARSPPLIC
FT
FT VARIANT 52 52
FT
FT VARIANT 52 52
FT
FT VARIANT 82 82
FT
FT VARIANT 199 199
FT
FT VARIANT 227 227
FT
FT CONFLICT 80 80
FT CONFLICT 85 85
FT CONFLICT 187 187
FT CONFLICT 297 297
SQ SEQUENCE 381 AA; 41388 MM; 29138EB6B4B565E CRC64;

Query Match 19.0%; Score 139; DB 1; Length 381;
Best Local Similarity 29.4%; Pred. No. 1.7e-06;
Matches 32; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

```

```

QY 19 GNFSSQCAFSCSEGTNLGIEETTC--GPRGNMSSPPTCQVIOCEPLSAPDLGIMN 75
DB 178 GILFGATISFCNTGKLGPTSLGLSSGVQWMDPLPECEIYCPAPQIDNIIQ 237
QY 76 CSPLASFSTACTFICSEGTLEIKKKKTTIC--ESSGIWSNPSPICQ 121
DB 238 GERD--HYGYROSIVTACNKGFTMIGHSITCYVNDGEMSGPPPCR 284

RESULT 40
DAF_PONPY STANDARD; PRT; 340 AA.
AC P49457;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE Complement decay-accelerating factor (CD55) (Fragment).
GN DAF OR CD55.
OS Pongo pygmaeus (Orangutan).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Pongo.
OX NCBI_TaxID=9600;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94110622; PubMed=7506731;
RA Nickells M.W., Alvarez J.I., Lublin D.M., Atkinson J.P.;
RT "Characterization of DAF-2, a high molecular weight form of decay-
RT accelerating factor (DAF, CD55), as a covalently cross-linked dimer
RT of DAF-1."
RL J. Immunol. 155:676-685(1994).
CC -1- FUNCTION: THIS PROTEIN RECOGNIZES C4B AND C3B FRAGMENTS THAT
CC CONDENSE WITH CELL-SURFACE HYDROXYL OR AMINO GROUPS WHEN NASCENT
CC C4B AND C3B ARE LOCALLY GENERATED DURING C4 AND C3 ACTIVATION.
CC INTERACTION OF DAF WITH CELL-ASSOCIATED C4B AND C3B POLYPEPTIDES
CC INTERFERES WITH THEIR ABILITY TO CATALYZE THE CONVERSION OF C2 AND
CC FACTOR B TO ENZYMATICALLY ACTIVE C2A AND BB AND THEREBY PREVENTS

```

```

CC THE FORMATION OF C4B2A AND C3BB, THE AMPLIFICATION CONVERTASES OF
CC THE COMPLEMENT CASCADE (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (MAJOR FORM) AND NON-DISULFIDE-LINKED, COVALENT
CC HOMODIMER (MINOR FORM).
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- ALTERNATIVE PRODUCTS: TWO FORMS OF DAF (DAF-2, SHOWN HERE, AND
CC DAF-1) ARE PRODUCED BY ALTERNATIVE SPLICING OF THE SAME GENE.
CC -1- FUNCTION: SCR2 AND SCR4 PROVIDE THE PROPER CONFORMATION FOR THE
CC ACTIVE SITE ON SCR3 (BY SIMILARITY).
CC -1- PTM: THE SER/THR-RICH DOMAIN IS HEAVILY O-GLYCOSYLATED.
CC -1- SIMILARITY: CONTAINS 4 SUSHI (SCR) DOMAINS.
CC -1- SIMILARITY: BELONGS TO THE RECEPTORS OF COMPLEMENT ACTIVATION
CC (RCA) FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; S67775; AAC60609.1; -.
DR HSSP; P08603; IHFI.
DR InterPro; IPR000436; Sush1_SCR_CCP.
DR Pfam; PF00084; sush1; 4.
DR SMART; SM00032; CCP; 4.
DR KMW Complement pathway; Plasma; Glycoprotein; Membrane; Repeat;
KW Alternative splicing; GPI-anchor; Sush1.
FT NON_TER 1
FT CHAIN <1 312
FT PROPEP 313 340
FT DOMAIN <1 54
FT DOMAIN <1 118
FT DOMAIN 121 180
FT DOMAIN 183 243
FT DOMAIN 246 315
FT DISULFID 24 53
FT DISULFID 57 104
FT DISULFID 88 117
FT DISULFID 122 163
FT DISULFID 149 179
FT DISULFID 184 226
FT DISULFID 212 242
FT CARBOHYD 54 54
FT CARBOHYD 107 107
FT LIPID 312 312
SQ SEQUENCE 340 AA; 37180 MM; D30865C058204290 CRC64;

Query Match 18.7%; Score 137; DB 1; Length 340;
Best Local Similarity 29.5%; Pred. No. 2.3e-06;
Matches 31; Conservative 16; Mismatches 50; Indels 8; Gaps 3;

```

```

QY 23 FSSQCAFSCSEGTNLGIEETTC--GPRGNMSSPPTCQVIOCEPLSAPDLGIMNCSHP 79
DB 141 FGATISFCNTGKLGPTSLGLSSGVQWMDPLPECEIYCPAPQIDNIIQGRND 200
QY 80 LASFSTACTFICSEGTLEIKKKKTTIC--ESSGIWSNPSPICQ 121
DB 201 --HYGYROSIVTACNKGFTMIGHSITCYVNDGEMSGPPPCR 243

```

Search completed: September 7, 2002, 10:23:46  
Job time: 539 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:15:28 : Search time 36.49 Seconds  
(without alignments)  
347.596 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328  
Perfect score: 733  
Sequence: 1 CEPLEAPDELGTMDCTHPEFGN.....WSNPSICQKDKSFSMIKE 132

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	727	99.2	385	1 A34015	L-selectin precurs
2	690	94.1	372	2 UC5377	L-selectin precurs
3	599	81.7	376	2 JC4892	L-selectin precurs
4	537	73.3	370	2 S22124	L-selectin precurs
5	535	73.0	372	2 S23936	L-selectin precurs
6	518	70.7	323	1 S09702	L-selectin precurs
7	518	70.7	372	1 A32375	L-selectin precurs
8	315	43.0	830	2 A30359	P-selectin precurs
9	313	42.7	768	2 I53821	P-selectin - rat
10	305.5	41.7	612	2 B42755	E-selectin precurs
11	300	40.9	768	2 A42755	P-selectin precurs
12	296	40.4	610	2 A35046	E-selectin precurs
13	280	39.6	646	2 JN0473	P-selectin precurs
14	288	39.3	551	2 I46709	P-selectin precurs
15	271.5	37.0	485	2 S36772	endothelial leukoc
16	245	33.4	482	2 JC5092	E-selectin - bovin
17	173	23.6	449	1 NBHDHS	E-selectin - pig
18	173	23.6	1231	1 NBHDH	complement factor
19	167.5	22.9	2014	2 I36936	complement factor
20	163.5	22.3	2489	2 I73012	complement recepto
21	160	21.8	597	1 S53711	complement C3b/C4b
22	155.5	21.2	1019	2 A38738	C4BP alpha chain P
23	155.5	20.5	560	2 T16833	complement factor
24	150.5	20.5	1025	1 A43526	hypothetical prote
25	150.5	20.5	1025	1 A43526	complement C3d/Eps
26	149.5	20.4	597	1 NBHUC4	C4b-binding protel
27	148	20.2	363	2 B45900	complement C3d/Eps
28	148	20.2	676	2 A45900	complement C3b rec
29	147.5	20.1	482	2 A34924	complement C3b/C4b

30	147.5	20.1	669	2 S65551	factor H - bovine
31	147.5	20.1	868	2 T20239	hypothetical prote
32	147.5	20.1	1927	2 T34288	hypothetical prote
33	146	19.9	558	2 S57953	C4BP protein alpha
34	145	19.8	610	1 I46001	C4b-binding protel
35	144.5	19.7	469	1 NBMSC4	C4BP protein beta
36	144	19.6	258	2 S57960	apolipoprotein H p
37	141.5	19.3	345	1 NBBO	complement C3d/Eps
38	141	19.2	1091	1 PL0009	complement factor
39	140.5	19.2	764	1 BBHO	apolipoprotein H h
40	139	19.0	263	1 WMVZSP	decay-accelerating
41	139	19.0	381	1 B26359	decay-accelerating
42	139	19.0	440	2 A26359	complement control
43	138.5	18.9	263	1 C36838	B18L protein - var
44	138.5	18.9	263	2 B72152	hypothetical prote
45	138.5	18.9	263	2 T28450	

## ALIGNMENTS

RESULT 1  
A34015  
L-selectin precursor, long splice form - human  
N:Alternate names: CD62L; leucocyte cell adhesion molecule-1 (LECAM-1, LAM-1); leukoc  
ral lymph node homing receptor Leu-8  
C:Species: Homo sapiens (man)  
C:date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: I55333; S06798; J10104; A34015; A33912  
R:Ord, D.C.; Ernst, T.J.; Zhou, L.J.; Rambaldi, A.; Sperlant, O.; Griffin, J.; Tedder  
J. Biol. Chem. 265, 7760-7767, 1990  
A:title: Structure of the gene encoding the human leukocyte adhesion molecule-1 (L1,  
A:Reference number: I55333; M01D:90243637  
A:Accession: I55333  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 14-385 <ORD>  
A:Cross-references: GB:M32414; NID:G187259; PIDN:AAB60700.1; PID:G386660  
R:Cammer, D.; James, S.P.; Stamenkovic, I.; Seed, B.  
Nature 342, 78-82, 1989  
A:title: Leu-8/T01 is the human equivalent of the Mel-14 lymph node homing receptor.  
A:Reference number: S06798; M01D:90044046  
A:Accession: S06798  
A:Molecule type: mRNA  
A:Residues: 1-225, 'S', 227-385 <CAM>  
A:Cross-references: EMBL:X17519; NID:G34344; PIDN:CAB43536.1; PID:G4902829  
R:Tedder, T.F.; Isaacs, C.M.; Ernst, T.J.; Demetri, G.D.; Adler, D.A.; Distche, C.M.  
J. Exp. Med. 170, 123-133, 1989  
A:title: Isolation and chromosomal localization of cDNAs encoding a novel human lymph  
nodeins.  
A:Reference number: J10104; M01D:89310350  
A:Accession: J10104  
A:Molecule type: mRNA  
A:Residues: 1-230, 'N', 232, 'N', 234-254, 'E', 256-385 <TRD>  
A:Cross-references: GB:X16150; NID:G34428; PIDN:CAA3275.1; PID:G34429  
A:Note: the translated sequence in Genbank entry HSLYAM1, release 111.0, differs from  
J. Bowen, B.R.; Nguyen, T.; Lasky, L.A.  
J. Cell Biol. 109, 421-427, 1989  
A:title: Characterization of a human homologue of the murine peripheral lymph node hc  
A:Reference number: A34015; M01D:89308881  
A:Accession: A34015  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 14-49, 'Y', 51-190, 'H', 192-205, 'V', 207-226, 'F', 228-385 <BOW>  
A:Cross-references: GB:X16070; NID:G38092; PIDN:CAA34203.1; PID:G38093  
R:Siegelman, M.H.; Weisman, I.L.  
Proc Natl. Acad. Sci. U.S.A. 86, 5562-5566, 1989  
A:title: Human homologue of mouse lymph node homing receptor: evolutionary conservati  
A:Reference number: A33912; M01D:89315837  
A:Accession: A33912  
A:Status: preliminary  
A:Molecule type: mRNA

A:Residues: 14-205; 'L', 207-385 <SIE>  
 A:Cross-references: GB:M25280; NID:9187182; PIDN:AA63053.1; PID:9307134  
 C:Comment: For an alternative splice form, see PIR:S09702.  
 C:Genetics:  
 A:Gene: GDB:SELL; GDB:LINR; LSEL; LAM1; LYAM1; LAM-1  
 A:Cross-references: GDB:120157; GDB:118834; OMIM:153240  
 A:Map position: 1q22-1q23  
 A:Introns: 14/3; 42/1; 171/1; 207/1; 269/1; 331/1; 374/1; 380/2  
 C:Function:  
 A:Description: binds with low affinity to oligosaccharides like heparan sulfate and sialyl ment of leucocytes to areas of inflammation, and with CD162 mediates neutrophil-neutroph C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology; C:Keywords: alternative splicing; cell adhesion; duplication; glycoprotein; inflammation  
 F:1-51/Domain: signal sequence #status predicted <SIG>  
 F:42-168/Domain: C-type lectin homology <LCH>  
 F:52-385/Product: L-selectin #status predicted <MAT>  
 F:52-343/Domain: extracellular #status predicted <EXT>  
 F:173-204/Domain: EGF homology <EGF>  
 F:210-267/Domain: complement factor H repeat homology <FH1>  
 F:272-329/Domain: complement factor H repeat homology <FH2>  
 F:344-368/Domain: transmembrane #status predicted <TM>  
 F:369-385/Domain: intracellular #status predicted <INT>  
 F:73,117,130,245,259/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:377,380/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 99.2%; Score 727; DB 1; Length 385;  
 Best Local Similarity 99.2%; Pred. No. 1,2e-57;  
 Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60  
 Db 210 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 269  
 Qy 61 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 120  
 Db 270 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 329  
 Qy 121 QKLDKSFMSIKE 132  
 Db 330 QKLDKSFMSIKE 341

RESULT 2  
 JC5377  
 L-selectin precursor - hamadryas baboon  
 C:Species: Papio hamadryas (hamadryas baboon)  
 C>Date: 02-Jun-1997 #sequence\_revision 12-Sep-1997 #text\_change 21-Jan-2000  
 C:Accession: JC5377; PC4315  
 R:Tsushita, N.; Fu, H.; Berg, E.L.  
 Gene 181, 219-220, 1996  
 A:Title: PCR cloning of the cDNA encoding baboon L-selectin.  
 A:Reference number: JC5377; MUID:97128794  
 A:Accession: JC5377  
 A:Molecule type: mRNA  
 A:Residues: 1-372 <TSU>  
 A:Cross-references: GB:U52074; NID:91326148; PIDN:AA640903.1; PID:91326149  
 A:Accession: PC4315  
 A:Molecule type: protein  
 A:Residues: 37-43;142-148 <TSU2>  
 C:Comment: This receptor is involved in the initial adhesive interaction between lymphoc sites of inflammation.  
 C:Superfamily: L-selectin; C-type lectin homology; complement factor H repeat homology; F:1-38/Domain: signal sequence #status predicted <SIG>  
 F:29-155/Domain: C-type lectin homology <LCH>  
 F:39-372/Product: L-selectin #status predicted <MAT>  
 F:39-157/Domain: calcium-binding #status predicted <CB>  
 F:160-191/Domain: EGF homology <EGF>  
 F:197-254/Domain: complement factor H repeat homology <FH1>  
 F:259-316/Domain: complement factor H repeat homology <FH2>  
 F:333-355/Domain: transmembrane #status predicted <TM>  
 F:356-372/Domain: intracellular #status predicted <INT>

Query Match 94.1%; Score 690; DB 2; Length 372;  
 Best Local Similarity 92.4%; Pred. No. 2.3e-54;  
 Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Qy 1 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60  
 Db 197 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 256  
 Qy 61 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 120  
 Db 257 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 316  
 Qy 121 QKLDKSFMSIKE 132  
 Db 317 QKLDKSFMSIKE 328

RESULT 3  
 JC4892  
 L-selectin precursor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C>Date: 04-Sep-1998 #sequence\_revision 04-Sep-1998 #text\_change 21-Jan-2000  
 C:Accession: JC4892  
 R:Qian, J.; Huang, X.; Marks, R.M.  
 Biochem. Biophys. Res. Commun. 225, 406-412, 1996  
 A:Title: Cloning of the cDNA for rabbit L-selectin and expression of recombinant prot A:Reference number: JC4892; MUID:96354800  
 A:Accession: JC4892  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-376 <OIA>  
 A:Cross-references: GB:U26535; NID:9847787; PIDN:AA67896.1; PID:9847788  
 C:Comment: This protein involved in leukocyte-endothelial adhesion; it mediates adhes F:1-37/Domain: signal sequence #status predicted <SIG>  
 F:29-155/Domain: C-type lectin homology <LCH>  
 F:38-376/Product: L-selectin #status predicted <MAT>  
 F:160-191/Domain: EGF homology <EGF>  
 F:197-254/Domain: complement factor H repeat homology <FH1>  
 F:259-316/Domain: complement factor H repeat homology <FH2>

Query Match 81.7%; Score 599; DB 2; Length 376;  
 Best Local Similarity 80.3%; Pred. No. 3.1e-46;  
 Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

Qy 1 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60  
 Db 197 CEPLAPLGTMDCTHPPGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 256  
 Qy 61 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 120  
 Db 257 IOCEPLAPDLGIMNCSHPPLASFSTACTFCISGTELGKKTTCSSGIMNSPSPIC 316  
 Qy 121 QKLDKSFMSIKE 132  
 Db 317 QKLDKSFMSIKE 328

RESULT 4  
 S22124  
 L-selectin precursor - bovine  
 N:Alternate names: leucocyte cell adhesion molecule 1 (LECAM-1)  
 C:Species: Bos primigenius taurus (cattle)  
 C>Date: 20-Feb-1995 #sequence\_revision 20-Feb-1995 #text\_change 21-Jan-2000  
 C:Accession: S22124; A46531  
 R:Bosworth, B.T.  
 submitted to the EMBL Data Library, October 1991  
 A:Reference number: S22123  
 A:Accession: S22124  
 A:Status: preliminary  
 A:Molecule type: mRNA













A:Reference number: A26505; MUID:83048213  
 A:Accession: A26505  
 A:Molecule type: protein  
 A:Residues: 19-20,'Q',22-29,'V',31-33,'Q',35 <SIM>  
 R:Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C. Biochemistry 31, 3626-3634, 1992  
 A:Title: Solution structure of the fifth repeat of factor H: A second example of the con  
 A:Reference number: A44551; MUID:92232649  
 R:Kristensen, T.; Wetzel, R.A.; Tack, B.F. J. Immunol. 136, 3407-3411, 1986  
 A:Title: Structural analysis of human complement protein H: homology with C4b binding pr  
 A:Reference number: S10479; MUID:86169701  
 A:Accession: S10479  
 A:Molecule type: mRNA  
 A:Residues: 226-401,'Y',403-449 <KRI>  
 A:Cross-references: GB:M12383; NID:g180472; PIDN:AAA52013.1; PID:g180473  
 C:Comment: Factor H has also been found bound to cell membranes in an unknown manner. HC  
 C:Comment: Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver.  
 C:Genetics: <HF1>  
 A:Gene: GDB:HF1; HF  
 A:Cross-references: GDB:120041; OMIM:134370  
 A:Map position: 1932-1932  
 C:Genetics: <HF2>  
 A:Gene: GDB:HF2; HF  
 A:Cross-references: GDB:129095  
 A:Map position: 1932-1932  
 A:Note: the correspondence between the two loci and the sequences indicated is unclear;  
 C:Function:  
 A:Description: a cofactor in the inactivation of C3b by serine proteinase I; also increas  
 he alternative complement pathway  
 A:Pathway: complement alternate pathway  
 C:Keywords: alternative splicing; complement alternate pathway; glycoprotein; plasma  
 F:118/Domain: signal sequence #status predicted <SIG>  
 F:19-449/Product: complement factor H, short splice form #status experimental <MAT>  
 F:21-80/Domain: complement factor H repeat homology <FH01>  
 F:85-141/Domain: complement factor H repeat homology <FH02>  
 F:146-205/Domain: complement factor H repeat homology <FH03>  
 F:210-262/Domain: complement factor H repeat homology <FH04>  
 F:246-248/Region: cell attachment (R-C-D) motif  
 F:267-320/Domain: complement factor H repeat homology <FH05>  
 F:325-385/Domain: complement factor H repeat homology <FH06>  
 F:389-442/Domain: complement factor H repeat homology <FH07>  
 F:21-66,52-80,85-129,114-141,146-192,178-205,210-251,237-262,267-309,294-320,325-374,357  
 F:217/Binding site: carbohydrate (Asn) (covalent) #status absent

A:Title: The complete amino acid sequence of human complement factor H.  
 A:Reference number: S00254; MUID:88134059  
 A:Accession: S00254  
 A:Molecule type: mRNA  
 A:Residues: 1-1231 <RIP>  
 A:Cross-references: EMBL:Y00716; NID:g31964; PIDN:CAA68704.1; PID:g31965  
 A:Note: 402-Tyr was also found  
 A:Note: parts of this sequence, including the amino and carboxyl ends of the mature p  
 R:Estaller, C.; Schwaebel, W.; Dierich, M.; Weiss, E.H. Eur. J. Immunol. 21, 799-802, 1991  
 A:Title: Human complement factor H: two factor H proteins are derived from alternativ  
 A:Reference number: A60238; MUID:91184292  
 A:Accession: A60238  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 1-56;1177-1231 <EST>  
 A:Note: only portions of this 4.3 kilobase mRNA were sequenced  
 R:Day, A.J.; Ripoché, J.; Lyons, A.; McIntosh, B.; Harris, T.J.R.; Sim, R.B. Biosci. Rep. 7, 201-207, 1987  
 A:Title: Sequence analysis of a cDNA clone encoding the C-terminal end of human compl  
 A:Reference number: A54726; MUID:88025472  
 A:Accession: A54726  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 'DFRN',579-1231 <DAY>  
 A:Cross-references: GB:M17517; NID:g180497; PIDN:AAA52016.1; PID:g180498  
 A:Note: parts of this sequence were determined by protein sequencing  
 R:Ripoché, J.; Day, A.J.; Willis, A.C.; Belt, K.T.; Campbell, R.D.; Sim, R.B. Biosci. Rep. 6, 65-72, 1986  
 A:Title: Partial characterization of human complement factor H by protein and cDNA se  
 A:Reference number: A61565; MUID:86186123  
 A:Accession: A61565  
 A:Status: not compared with conceptual translation  
 A:Molecule type: mRNA  
 A:Residues: 'METSNGHNAKT',1050-1057,'T',1059-1102 <RU2>  
 R:Sim, R.B.; Discipio, R.G. Biochem. J. 205, 285-293, 1982  
 A:Title: Purification and structural studies on the complement-system control protein  
 A:Reference number: A26505; MUID:83048213  
 A:Accession: A26505  
 A:Molecule type: protein  
 A:Residues: 19-20,'Q',22-29,'V',31-33,'Q',35 <SIM>  
 R:Barlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P Biochemistry 31, 3626-3634, 1992  
 A:Title: Solution structure of the fifth repeat of factor H: A second example of the  
 A:Reference number: A44551; MUID:92232649  
 R:Kristensen, T.; Wetzel, R.A.; Tack, B.F. J. Immunol. 136, 3407-3411, 1986  
 A:Title: Structural analysis of human complement protein H: homology with C4b binding pr  
 A:Reference number: S10479; MUID:86169701  
 A:Accession: S10479  
 A:Molecule type: mRNA  
 A:Residues: 226-401,'Y',403-449 <KRI>  
 A:Cross-references: GB:M12383; NID:g180472; PIDN:AAA52013.1; PID:g180473  
 C:Comment: Factor H has also been found bound to cell membranes in an unknown manner. HC  
 C:Comment: Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liv  
 C:Genetics: <HF1>

Query Match 23.6%; Score 173; DB 1; Length 449;  
 Best Local Similarity 28.8%; Pred. No. 4.1e-08;  
 Matches 36; Conservative 25; Mismatches 48; Indels 16; Gaps 5;  
 QY 14 CTH-----PFGNFS-----FSSQCAFSCEGNTLNG-IEETTCGPRGNSSPPPTQGV 60  
 DB 85 CGHGGRDPFGFTLLGNGVFEYGVKAVYTCNEGYQLGEIYRCDYDVG-WTNDPIPEV 143  
 QY 61 IQCEPLASPDG--IMNSHPIASFSPACTFICSEGTLEIGKKTKICSSGSIANSNP 118  
 DB 144 VKCLPVARPEKGIYSSAMEPDREYHFGQAVRFVPCNSYKIEGDEHMCSDGFEWSEKRP 203  
 QY 119 ICOKL 123  
 DB 204 KCVEI 208

RESULT 18  
 NBH0H  
 complement factor H precursor, long splice form [validated] - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 08-Dec-2000  
 C:Accession: S00254; A60238; A54726; A61565; A26505; 172654; S66298  
 R:Ripoché, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.  
 Biochem. J. 249, 593-602, 1988

Query Match 23.6%; Score 173; DB 1; Length 449;  
 Best Local Similarity 28.8%; Pred. No. 4.1e-08;  
 Matches 36; Conservative 25; Mismatches 48; Indels 16; Gaps 5;  
 QY 14 CTH-----PFGNFS-----FSSQCAFSCEGNTLNG-IEETTCGPRGNSSPPPTQGV 60  
 DB 85 CGHGGRDPFGFTLLGNGVFEYGVKAVYTCNEGYQLGEIYRCDYDVG-WTNDPIPEV 143  
 QY 61 IQCEPLASPDG--IMNSHPIASFSPACTFICSEGTLEIGKKTKICSSGSIANSNP 118  
 DB 144 VKCLPVARPEKGIYSSAMEPDREYHFGQAVRFVPCNSYKIEGDEHMCSDGFEWSEKRP 203  
 QY 119 ICOKL 123  
 DB 204 KCVEI 208

RESULT 18  
 NBH0H  
 complement factor H precursor, long splice form [validated] - human  
 C:Species: Homo sapiens (man)  
 C:Date: 31-Dec-1993 #sequence revision 31-Dec-1993 #text change 08-Dec-2000  
 C:Accession: S00254; A60238; A54726; A61565; A26505; 172654; S66298  
 R:Ripoché, J.; Day, A.J.; Harris, T.J.R.; Sim, R.B.  
 Biochem. J. 249, 593-602, 1988

```

A:Status: preliminary; translated from GB/EMBL/DBD1
A:Molecule type: mRNA
A:Residues: 1-2014 <BIR1>
A:Cross-references: GB:I24920; NID:g551564; PIDN:AAA51438.1; PID:g557725
A:Accession: I56937
A:Status: preliminary; translated from GB/EMBL/DBD1
A:Molecule type: mRNA
A:Residues: 1-397,1751-2014 <BIR2>
A:Cross-references: GB:I24921; NID:g557726; PIDN:AAA51439.1; PID:g557727
A:Genetics:
A:Gene: CRI
A:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology
C:Keywords: duplication; glycoprotein; tandem repeat
F:18-74/Domain: complement factor H repeat homology <FH01>
F:79-136/Domain: complement factor H repeat homology <FH02>
F:141-207/Domain: complement factor H repeat homology <FH03>
F:213-268/Domain: complement factor H repeat homology <FH04>
F:227-328/Domain: complement factor H repeat homology <FH05>
F:333-351/Domain: complement factor H repeat homology <FH06>
F:396-462/Domain: complement factor H repeat homology <FH07>
F:470-526/Domain: complement factor H repeat homology <FH08>
F:529-586/Domain: complement factor H repeat homology <FH09>
F:722-778/Domain: complement factor H repeat homology <FH10>
F:1041-1107/Domain: complement factor H repeat homology <FH11>
F:1172-1228/Domain: complement factor H repeat homology <FH12>
F:1233-1291/Domain: complement factor H repeat homology <FH13>
F:1296-1362/Domain: complement factor H repeat homology <FH14>
F:1432-1489/Domain: complement factor H repeat homology <FH15>
F:1625-1681/Domain: complement factor H repeat homology <FH16>
F:1749-1815/Domain: complement factor H repeat homology <FH17>
F:1823-1879/Domain: complement factor H repeat homology <FH18>
F:1884-1940/Domain: complement factor H repeat homology <FH19>

Query Match 22.9%; Score 167.5; DB 2; Length 2014;
Best Local Similarity 32.6%; Pred. No. 5.5e-07;
Matches 46; Conservative 16; Mismatches 56; Indels 23; Gaps 6;

Oy 1 CEPLAPELGIMDCNHPGNFSSQCAFSCSECTNLGIEETGCGFGN---WSSEPT 57
Db 1432 CGPPEPPNGAV---HMTDPOFGSTVYNSCNEFRLGSRSTTCLVSGNNVTDKKAPI 1488
Oy 58 TQVIOCEP---LSAPDLGIMNCSPHLASFSTSACTFCSEGT-----ELICKRTICE 108
Db 1489 CEIICEPPEPTISNDF---YSNNRASFGNGTAVTQCHTGPGEOLFELVGERISYCT 1544
Oy 109 SS-----GIWSNPSPLCQKLK 125
Db 1545 SKDDVGVWSSPPRCISTNK 1565

RESULT 20
T73012
complement C3b/C4b receptor, membrane-bound form precursor - human
N:Alternate names: complement C3b/C4b receptor; complement receptor type 1 (CRI); sur
N:Contains: complement C3b/C4b receptor, secreted form
C:Species: Homo sapiens (man)
C:Date: 02-Jul-1996 #sequence,revision 24-Nov-1999 #text,change 21-Jul-2000
C:Accession: I73012; I56203; A47602; S03291; S03843; A28507; A24748; C24748
R:VLX, D.P.; Mong, W.W.
R:Immunol. 151, 6214-6224, 1993
A:Title: Structure of the gene for the F allele of complement receptor type 1 and seq
A:Reference number: I56203; MUID:9406175
A:Accession: I73012
A:Status: preliminary; translated from GB/EMBL/DBD1
A:Molecule type: DNA
A:Residues: 1-683,'X',685-894,'A',896-1000,1451-1471,'X',1473-2489 <VIK1>
A:Cross-references: GB:LI7418; NID:g306678; PIDN:AAB60694.1; PID:g306680
A:Accession: I56203
A:Status: translated from GB/EMBL/DBD1
A:Molecule type: DNA
A:Residues: 1-683,'X',685-894,'A',896-1000,1451-1471,'X',1473-2489 <VIK2>
A:Cross-references: GB:LI7418; NID:g306678; PIDN:AAB60694.1; PID:g306680

```





Db 184 KHNGGNEIDHTYGVSSVYSCNPFSLGSEASISCTVKNKTVGVWSPSPVCKEI 237

RESULT 22

MEMSH

complement factor H precursor - mouse

N:Alternate names: protein beta-1-H

C:Species: Mus musculus (house mouse)

C>Date: 30-Sep-1987 #sequence\_revision 30-Sep-1987 #text\_change 22-Jun-1999

C:Accession: A26154; 149711; 149728

R:Kristensen, T.; Tack, B.F.

Proc. Natl. Acad. Sci. U.S.A. 83, 3963-3967, 1986

A:Title: Murine protein H is comprised of 20 repeating units, 61 amino acids in length.

A:Reference number: A26154; MUID:86233353

A:Accession: A26154

A:Molecule type: mRNA

A:Residues: 1-1234 <KR1>

A:Cross-references: GB:M12660; NID:9193724; PIDN:AAA37759.1; PID:9387181

R:Natsuume-Sakai, S.; Nonaka, M.; Nonaka, M.; Harada, Y.

J. Immunol. 144, 358-362, 1990

A:Title: Demonstration of an unusual allelic variation of mouse factor H by the complete

A:Reference number: 149711; MUID:90111033

A:Accession: 149711

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-18 <RES>

A:Cross-references: GB:M31979; NID:9193726; PIDN:AAA37762.1; PID:9193729

R:Munoz-Canoves, P.; Tack, B.F.; Vlk, D.P.

Biochemistry 28, 9891-9897, 1989

A:Title: Analysis of complement factor H mRNA expression: Dexamethasone and IFN-gamma in

A:Reference number: 149728; MUID:90148935

A:Accession: 149728

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-19 <RE2>

A:Cross-references: GB:J02891; NID:9193805; PIDN:AAA37795.1; PID:9553926

C:Comment: Two codominant alleles of factor H are present in mice.

C:Comment: Factor H functions as a cofactor in the inactivation of C3b by serine proteinase (C5 convertase) in the alternative complement pathway.

C:Genetics:

A:Map position: 1

C:Superfamily: complement factor H; complement factor H repeat homology

C:Keywords: complement alternate pathway; duplication; glycoprotein; plasma

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-123/Domain: complement factor H #status predicted <MPT>

F:21-80/Domain: complement factor H repeat homology <FH01>

F:85-141/Domain: complement factor H repeat homology <FH02>

F:146-205/Domain: complement factor H repeat homology <FH03>

F:210-262/Domain: complement factor H repeat homology <FH04>

F:246-248/Region: cell attachment (R-G-D) motif

F:267-320/Domain: complement factor H repeat homology <FH05>

F:325-385/Domain: complement factor H repeat homology <FH06>

F:389-442/Domain: complement factor H repeat homology <FH07>

F:448-505/Domain: complement factor H repeat homology <FH08>

F:509-564/Domain: complement factor H repeat homology <FH09>

F:569-622/Domain: complement factor H repeat homology <FH10>

F:620-683/Domain: complement factor H repeat homology <FH11>

F:690-743/Domain: complement factor H repeat homology <FH12>

F:752-802/Domain: complement factor H repeat homology <FH13>

F:808-861/Domain: complement factor H repeat homology <FH14>

F:867-931/Domain: complement factor H repeat homology <FH15>

F:936-989/Domain: complement factor H repeat homology <FH16>

F:994-1048/Domain: complement factor H repeat homology <FH17>

F:1053-1107/Domain: complement factor H repeat homology <FH18>

F:1114-1168/Domain: complement factor H repeat homology <FH19>

F:1172-1233/Domain: complement factor H repeat homology <FH20>

F:21-66,52-80,85-129,114-141,146-192,178-205,210-251,237-262,267-309,294-320,325-374,357-402,808-850,836-861,867-920,906-931,936-978,964-989,994-1037,1023-1048,1053-1096,1082-1676,721,773,801,1030,1061,1225/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 21.8%; Score 160; DB 1; Length 1234;

Best Local Similarity 28.0%; Pred. No. 1.6e-06;

Matches 35; Conservative 25; Mismatches 49; Indels 16; Gaps 5;

QY 14 CTH-----PFCN-----FSFSSQCAFSCSEBNTLTC-IEFTCCGFGWMSPEPTCOY 60

Db 85 CGHPGTPFGSFRPLAVGSGFEAGKVVYTCDDYQLLGEIDREGCADG-WINDPLCEV 143

QY 61 IOCEPLSAPDLG--INMCSHPLASFSTSCACRFTICSEGTLLCKKRTICSSGNSPSP 118

Db 144 VKCLPYTELENGRIVGGAETDDEYFGQVYVRECSGFRTEBKHKEHCSENLMSNEKP 203

QY 119 ICQKL 123

Db 204 RCVEI 208

RESULT 23

A38738

coagulation factor C precursor - horseshoe crab (Tachyplesus tridentatus)

N:Alternate names: coagulation-complement factor C; Limulus factor C

N:Contains: coagulation factor C heavy chain; coagulation factor C light chain peptid

C:Species: Tachyplesus tridentatus

C>Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 08-Dec-2000

C:Accession: A38738; B38738; S00105

R:Muta, T.; Miyata, T.; Mism, Y.; Tokunaga, F.; Nakamura, T.; Toh, Y.; Ikehara, Y.;

J. Biol. Chem. 266, 6554-6561, 1991

A:Title: Limulus factor C. An endotoxin-sensitive serine protease zymogen with a mosa

A:Reference number: A38738; MUID:91177916

A:Accession: A38738

A:Molecule type: mRNA

A:Residues: 1-1019 <MUT>

A:Cross-references: GB:D90271; NID:9217396; PID:0105020; PID:9217397

A:Accession: B38738

A:Molecule type: mRNA

A:Residues: 1-466,616, 'DN', 619-620, 'A', 622 <MUT2>

A:Cross-references: GB:D90272

R:Tokunaga, F.; Miyata, T.; Nakamura, T.; Morita, T.; Kuma, K.I.; Miyata, T.; Iwanaga

Eur. J. Biochem. 167, 405-416, 1987

A:Title: Lipopolysaccharide-sensitive serine-protease zymogen (factor C) of horseshoe

is a novel type of serine protease

A:Reference number: S00105; MUID:88004461

A:Accession: S00105

A:Molecule type: protein

A:Residues: 'S', 27-39, 'S', 41, 'T', 43-54, 'E', 691-782,950-977 <TOR>

C:Superfamily: coagulation factor C; C-type lectin homology; complement factor H repe

C:Keywords: alternative splicing; glycoprotein; hemolymph coagulation; hydrolase; ser

F:1-25/Domain: signal sequence #status predicted <SIG>

F:26-690/Product: coagulation factor C heavy chain #status experimental <HCH>

F:136-195/Domain: complement factor H repeat homology <FH01>

F:199-254/Domain: complement factor H repeat homology <FH02>

F:260-321/Domain: complement factor H repeat homology <FH03>

F:436-564/Domain: C-type lectin homology <LCH>

F:576-634/Domain: complement factor H repeat homology <FH04>

F:685-747/Domain: complement factor H repeat homology #status atypical <FH05>

F:691-762/Product: coagulation factor C light chain peptide A #status experimental <P

F:763-1019/Product: coagulation factor C light chain peptide B #status experimental <

F:763-1015/Domain: trypsin homology <TRY>

F:523,534,624,912/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:767/Binding site: carbohydrate (Asn) (covalent) #status experimental

F:809,865,966/Active site: His, Asp, Ser #status predicted

Query Match 21.2%; Score 155.5; DB 2; Length 1019;

Best Local Similarity 34.7%; Pred. No. 3.4e-06;

Matches 33; Conservative 16; Mismatches 43; Indels 3; Gaps 2;

QY 29 FSCSEGTNLGIEETTCGFGWMSPEPTCOVYQCEPLSNAPDLGINMCSHPLASFSTSA 88

Db 166 YSCSPGFKLGVARISCLPFGWMSPPRC-IRECAKVSPEHGKVN--PSCNMTGAT 222

QY 89 CTFCEGHELGKKTICSSGINSNPFICOKL 123

Db 223 LRFSCDSPYLLIGQETLTCGNGWMSGQIPQCKKL 257



A:Title: Derivation of the sequence of the signal peptide in human C4b-binding protein A  
A:Reference number: S02372; MUID:88242821  
A:Accession: S02372  
A:Molecule type: mRNA  
A:Residues: 17-81 <L12>  
A:Cross-references: EMBL:X07853  
A:Note: although the sequence determined extends to residue 9 above, these authors assign  
R.Chung, L.P.; Bentley, D.R.; Reid, K.B.M.  
Biochem. J. 230, 133-141, 1985  
A:Title: Molecular cloning and characterization of the cDNA coding for C4b-binding prote  
A:Reference number: A90326; MUID:86025405  
A:Accession: A90326  
A:Molecule type: mRNA  
A:Residues: 80-597 <CH2>  
A:Cross-references: GB:X02865; NID:g29564; PIDN:CAA26617.1; PID:g29565  
A:Note: 92-Thr and 357-His were also found  
R.Lintin, S.J.; Reid, K.B.M.  
FEBS Lett. 204, 77-81, 1986  
A:Title: Studies on the structure of the human C4b-binding protein gene.  
A:Reference number: A24182; MUID:86301119  
A:Accession: A24182  
A:Molecule type: DNA  
A:Residues: 203-288 <LIN>  
A:Cross-references: EMBL:X04284; EMBL:X04296  
R.Rodriguez de Cordoba, S.; Sanchez-Corral, P.; Rey-Campos, J.  
J. Exp. Med. 173, 1073-1082, 1991  
A:Title: Structure of the gene coding for the alpha polypeptide chain of the human comp  
A:Reference number: A43023; MUID:91217619  
A:Contents: annotation; exon-intron boundaries  
R.Chung, L.P.; Gagnon, J.; Reid, K.B.M.  
Mol. Immunol. 22, 427-435, 1985  
A:Title: Amino acid sequence studies of human C4b-binding protein: N-terminal sequence a  
ced by cyanogen bromide treatment.  
A:Reference number: A93134; MUID:85296001  
A:Accession: A93134  
A:Molecule type: protein  
A:Residues: 49-81 <CH1>  
A:Note: this paper reports amino-terminal sequences of the intact protein and of a numbe  
R.Hessling, M.; Kanters, D.; Takeya, H.; van't Veer, C.; Hackeng, T.M.; Iwanaga, S.; Boum  
FEBS Lett. 317, 228-232, 1993  
A:Title: The region Ser(33)-Arg(356) of the alpha-chain of human C4b-binding protein is  
A:Reference number: S29492; MUID:93146164  
A:Accession: S29492  
A>Status: preliminary  
A:Molecule type: protein  
A:Residues: 381-404 <HES>  
R.Suzuki, K.; Nishio, J.  
J. Biol. Chem. 263, 17034-17039, 1988  
A:Title: Binding site for vitamin K-dependent protein S on complement C4b-binding protei  
A:Reference number: A31785; MUID:89034204  
A:Accession: A31785  
A:Molecule type: protein  
A:Residues: 495-505, 'X', 507-510, 'X', 512-515 <SUZ>  
A:Note: this peptide appears to bind protein S  
R.Dahlback, B.; Smith, C.A.; Muller-Eberhard, H.J.  
Proc. Natl. Acad. Sci. U.S.A. 80, 3461-3465, 1983  
A:Title: Visualization of human C4b-binding protein and its complexes with vitamin K-def  
A:Reference number: A93950; MUID:83221615  
A:Contents: annotation; electron microscopy; three-dimensional structure; ligand binding  
R.Aso, T.; Okamura, S.; Matsuguchi, T.; Sakamoto, N.; Sata, T.; Niho, Y.  
Biochem. Biophys. Res. Commun. 174, 222-227, 1991  
A:Title: Genomic organization of the alpha chain of the human C4b-binding protein gene.  
A:Reference number: I52244; MUID:91113199  
A:Accession: I52244  
A:Molecule type: DNA  
A>Status: translation not shown; translated from GB/EMBL/DBJ  
A:Residues: 1-597 <ASO>  
A:Cross-references: GB:M62486; NID:g190498; PIDN:AAA36506.1; PID:g190500  
C:Comment: C4BP controls the classical pathway of complement activation. It binds as a c  
the degradation of the C4bC2a complex (C3 convertase) by dissociating the complement fr  
C:Comment: C4BP occurs in plasma in two forms, both of which bind complement fragment C4  
sulfilde bonds. The beta chain binds the vitamin K-dependent plasma protein S. A minor f  
actors V and VIII.

C:Comment: The molecule has a central body supporting seven tentacles (alpha chains),  
C:Genetics:  
A:Gene: GDB:C4BPA  
A:Cross-references: GDB:120568; OMTM:120830  
A:Map position: 1q32-1q32  
A:Map position: 1q32-1q32  
A:Introns: 48/1; 110/1; 143/2; 172/1; 236/1; 297/1; 362/1; 425/1; 482/1; 540/3  
C:Complex: octamer of seven alpha chains and one beta chain  
C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology  
C:Keywords: acute phase; chylomicron; complement pathway; duplication; glycoprotein;  
F:1-48/Domain: signal sequence #status predicted <SIG>  
F:49-597/Product: C4b-binding protein alpha chain #status predicted <MAT>  
F:50-108/Domain: complement factor H repeat homology <FH1>  
F:113-170/Domain: complement factor H repeat homology <FH2>  
F:175-234/Domain: complement factor H repeat homology <FH3>  
F:239-294/Domain: complement factor H repeat homology <FH4>  
F:299-360/Domain: complement factor H repeat homology <FH5>  
F:364-422/Domain: complement factor H repeat homology <FH6>  
F:381-404/Region: complement C4b binding #status predicted  
F:426-480/Domain: complement factor H repeat homology <FH7>  
F:484-538/Domain: complement factor H repeat homology <FH8>  
F:521,506,528/Binding site: carbohydrate (Asn) (covalent) #status experimental  
Query Match 20.4%; Score 149.5; DB 1; Length 597;  
Best Local Similarity 29.1%; Pred. No. 6.8e-06;  
Matches 37; Conservative 18; Mismatches 53; Indels 19; Gaps 4;  
QY 14 CTHP-----FGNFSFSSQCAFSCSEGTNLGIEETTCGPRG---NWSSEPTTCOV 60  
DB 113 CHNPGELRMNGVEIKTDSFGSIEFSCSEGFLLGSTRSCEVDGRGVGWHPLPQCEI 172  
QY 61 ICCEPLASAPDLICMNCSPHLASFTFSACTFTICSETELGKKKTYC---ESSGIWNP 116  
DB 173 VKCKR--PPDINRGRSGEENFYAVGFVSYSCDPRFLGLHASISCTVENETIGWRPS 230  
QY 117 SPICOKL 123  
DB 231 PPTCEKI 237  
RESULT 27  
B45900  
complement C3d/Epstein-Barr virus receptor precursor alternative splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
C:Accession: B45900; C45900  
R.Kurtz, C.B.; O'Toole, E.; Christensen, S.M.; Wels, J.H.  
J. Immunol. 144, 3581-3591, 1990  
A:Title: The murine complement receptor gene family. IV. Alternative splicing of Cr2  
A:Reference number: A45900; MUID:90229754  
A:Accession: B45900  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-363 <KUR>  
A:Accession: C45900  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual  
A:Molecule type: mRNA  
A:Residues: 1-19, 'E', 84-363 <KU2>  
C:Superfamily: complement C3d/Epstein-Barr virus receptor; complement factor H repeat  
F:23-81/Domain: complement factor H repeat homology <FH1>  
F:86-145/Domain: complement factor H repeat homology <FH2>  
F:154-210/Domain: complement factor H repeat homology <FH3>  
F:218-274/Domain: complement factor H repeat homology <FH4>  
F:279-335/Domain: complement factor H repeat homology <FH5>  
Query Match 20.2%; Score 148; DB 2; Length 363;  
Best Local Similarity 32.1%; Pred. No. 5.7e-06;  
Matches 43; Conservative 17; Mismatches 58; Indels 16; Gaps 7;  
QY 1 CE---PLEAPELGTMDCTHPG---NFSFSSQCAFSCSEGTNLGIEETTCGPRGNSS 53  
DB 210 CESDPELCEPSPPTTHNGHHTGQHYDVAGLVSYTSCPEGFLTGKTKIKCLSGSDMDG 269

[illegible]

RESULT 28

A45900  
complement C3b receptor type 2 long form precursor - mouse (fragment)

C:Species: Mus musculus (house mouse)  
C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 08-Oct-1999  
C:Accession: A45900; I48306  
R:Kurtz, C.B.; O'Toole, E.; Christensen, S.M.; Weis, J.H.  
J:Immunol. 144, 3581-3591, 1990  
A:Title: The murine complement receptor gene family. IV. Alternative splicing of Cr2 gene  
A:Reference number: A45900; MUID:90229754  
A:Accession: A45900  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-676 <KUR>  
A:Cross-references: GB:M36470  
A:Experimental source: clone 3i-1  
R:Kim, Y.U.; Kinoshita, T.; Molina, H.; Hourcade, D.; Seya, T.; Wagner, L.M.; Holers, V.  
J: Exp. Med. 181, 151-159, 1995  
A:Title: Mouse complement regulatory protein Cr1/p65 uses the specific mechanisms of bc  
A:Reference number: I48306; MUID:95105691  
A:Accession: I48306  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 21-367 <RES>  
A:Cross-references: EMBL:U17128; NID:g595980; PIDN:AA78271.1; PID:g595982  
C:Genetics:  
A:Gene: Cr2  
A:Introns: 80/1; 113/2; 142/1; 275/1; 333/1  
C:Superfamily: complement factor H repeat homology  
F:22-78/Domain: complement factor H repeat homology <FH01>  
F:83-140/Domain: complement factor H repeat homology <FH02>  
F:145-211/Domain: complement factor H repeat homology <FH03>  
F:217-272/Domain: complement factor H repeat homology <FH04>  
F:276-331/Domain: complement factor H repeat homology <FH05>  
F:336-394/Domain: complement factor H repeat homology <FH06>  
F:399-458/Domain: complement factor H repeat homology <FH07>  
F:467-523/Domain: complement factor H repeat homology <FH08>  
F:531-587/Domain: complement factor H repeat homology <FH09>  
F:592-648/Domain: complement factor H repeat homology <FH10>

```

Query Match          20.2%; Score 148; DB 2; Length 676;
Best Local Similarity 26.48; Pred. No. 1,1e-05;
Matches 39; Conservative 24; Mismatches 51; Indels 34; Gaps

QY 1 CEPEE-APEIGT-MDCTAP-----FGN-----FSFSQCAFSCSEGTNLGIETTC 45
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 259 CQPMKSKSPPLPCFMOCLEPQNLHGDYNNKKEFFSVQGNKYTCNPBYTLIGNIVEC 318
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 46 GPFSNMSSPEPTCOVIOCEPLSAPDLGIMNCNHPL-----ASFSTISACTFTSSEGT 98
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 319 TSLTQSWNTNPTCEVKKSCDAIP-----NHLHGRVFLPPNLQLGAEVSEVDLDGQ 369
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 99 LIGKKITICSSG---IWSNPSPICOKL 123
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 370 LKGRPSSOCIPEGETYIMNKKFPVCEOI 397
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 29
A34924
complement C3b/Cab receptor-like protein precursor - human
C:Species: Homo sapiens (man)
#:Date: 27-Jul-1990 #sequence_revision 11-Apr-1997 #text_change 03-Dec-1999

```

C:Accession: D044924; S03292  
R:Hourcade, D.; Miesner, D.R.; Bee, C.; Zeldes, W.; Atkinson, J.P.  
J. Biol. Chem. 265, 974-980, 1990  
A:Title: Duplication and divergence of the amino-terminal coding region of the comple  
A:Reference number: A34924; MUID:90110163  
A:Accession: A34924  
A:Molecule type: DNA  
A:Residues: 1-479 <HOU>  
A:Cross-references: GB:J05195  
R:Hourcade, D.; Miesner, D.R.; Atkinson, J.P.; Holers, V.M.  
J. Exp. Med. 168, 1255-1270, 1988  
A:Title: Identification of an alternative polyadenylation site in the human C3b/C4b r  
type 1.  
A:Reference number: S03291; MUID:89010527  
A:Accession: S03292  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: FPFYAFR, 452-482 <HO2>  
A:Cross-references: EMBL:X14360  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1988  
C:Accession: C4b-binding protein alpha chain; complement factor H repeat homology  
F:43-99/Domain: complement factor H repeat homology <FH01>  
F:104-161/Domain: complement factor H repeat homology <FH02>  
F:166-232/Domain: complement factor H repeat homology <FH03>  
F:238-293/Domain: complement factor H repeat homology <FH04>  
F:297-353/Domain: complement factor H repeat homology <FH05>  
F:358-416/Domain: complement factor H repeat homology <FH06>  
F:421-478/Domain: complement factor H repeat homology <FH07>

Query	March	20.1%	Score 147.5	DB 2:	length 482;
	Best Local Similarity	29.6%	Pred. No. 8.4e-06;		
	Matches	32;	Conservative 19;	Mismatches 48;	Indels 9; Gaps 3.
QY	20	NFSFSGCAFCSCSECTNLGITETTCGPGMNSSPDEPQCIVQCEPLSAPDGIIMNC	SHP	79	
Db	315	NFSFGQEFVYFSCPEYDLDGSGTYLHCTPGDGMSPAPRCVASCDF	---	LGQLPNGVH	370
QY	80	L--ASFSTSACTFCSEGTETLIGKKRTICSSG---	IWSNBPICOK	122	
Db	311	LFPNLQIGAKYDFVCDSEFGFKGSSASTCVYLAGHESLNNSSVPCEK		418	

```

RESULT      30
S65551
factor H - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 20-Jun-2000
C:Accession: S65551
R:Soames, C.J.; Day, A.J.; Sim, R.B.
Biochem. J. 315, 523-531, 1996
A:Title: Prediction from sequence comparisons of residues of factor H involved in the
A:Reference number: S65551; MUID:96202005
A:Accession: S65551
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-669 <SOA>
A:Cross-references: GB:V98697; NID:g1419423; PIDN:CA67257.1; PID:g1419424
C:Superfamily: complement factor H; complement factor H repeat homology <FHR1>
F:55-114/Domain: complement factor H repeat homology <FHR2>
F:256-349/Domain: complement factor H repeat homology <FHR3>
F:355-412/Domain: complement factor H repeat homology <FHR4>
F:416-471/Domain: complement factor H repeat homology <FHR5>
F:476-530/Domain: complement factor H repeat homology <FHR6>
F:538-582/Domain: complement factor H repeat homology <FHR5>
F:599-651/Domain: complement factor H repeat homology <FHR6>

```

Query Match	20.1%;	Score 147.5;	DB 2;	Length 669;
Best Local Similarity	27.5%;	Pred. No. 1.2e-05;		
Matches 30;	Conservative 26;	Mismatches 48;	Indels 5;	Gaps 4;
19 GN-PSFSQCAFSCSEGNLNG-IDETTCGPGGNNSSPEPTQVITQCEPLSADLG--IM 74				

Db 10 GNGEYGAKKVYTCDEGVQWVGEMNFRECDTNG--WTNDIPICEVYKCLPVTPEENGKIFS 68  
 Oy 75 NCSHPLASFSTTSACTFICSGTELIGKKTKITCESSGIWNSPPICOKL 123  
 Db 69 DALEPDEYTYGOVVOFECSNGYMLDGPQIHCISAGVWSAETPKCVEI 117

## RESULT 31

T20239  
 hypothetical protein C54G4.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
 C:Accession: T20239  
 R:Millerson, J.  
 submitted to the EMBL Data Library, June 1996  
 A:Reference number: 219241  
 A:Accession: T20239  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-868 <MIL>  
 A:Cross-references: EMBL:Z75533; PTDN:CAA99822.1; GSPDB:GN00019; CESP:C54G4.4  
 A:Experimental source: clone C54G4  
 C:Genetics:  
 A:Gene: CESP:C54G4.4  
 A:Map position: 1  
 A:Introns: 189/2; 254/2; 291/3; 339/1; 380/2; 536/1; 562/3; 661/1; 756/3; 829/1

Query Match 20.1%; Score 147.5; DB 2; Length 868;  
 Best Local Similarity 31.4%; Pred. No. 1.5e-05;  
 Matches 33; Conservative 14; Mismatches 47; Indels 11; Gaps 3;

Oy 23 FSSQCAFSCSEGNITGIEET-----CGPFGNWSPEPTCOYIQCEPLSAPDLGIMN 75  
 Db 559 FESEVEYECNSGMHLANSPSYSRLRRVCSQSDGIMSGSEPTCELVDC---GRPL-IAN 614  
 Oy 76 CSHPLASFSTTSACTFICSGTELIGKKTKITCESSGIWNSPPICOKL 120  
 Db 615 GRVDESSTFESANVYTCGQFRLIGPESLMCGDRGEMQPATPFC 659

## RESULT 32

T34288  
 hypothetical protein F47C12.1 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T34288  
 R:Murray, J.; Mohldmann, P.; Sansone, J.  
 submitted to the EMBL Data Library, June 1996  
 A:Description: The sequence of C. elegans cosmid F47C12.  
 A:Reference number: 221499  
 A:Accession: T34288  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1827 <MUR>  
 A:Cross-references: EMBL:U61946; PTDN:AAC24388.1; GSPDB:GN00022; CESP:F47C12.1  
 C:Experimental source: Strain Bristol N2; clone F47C12  
 C:Genetics:  
 A:Gene: CESP:F47C12.1  
 A:Map position: 4  
 A:Introns: 66/2; 221/1; 282/1; 404/1; 500/1; 649/3; 750/1; 997/2; 1152/3; 1279/1; 1346/1

Query Match 20.1%; Score 147.5; DB 2; Length 1827;  
 Best Local Similarity 27.1%; Pred. No. 3.1e-05;  
 Matches 39; Conservative 18; Mismatches 48; Indels 39; Gaps 7;

Oy 1 CEPL---EAPL---GTMDCTHFGNFSFSSQCAFSCSBS---TNLTGIEETTCGPPGNMS 52  
 Db 280 COPICGAVPOJANGFVDSA---TNVSTCGQVYKSCHKGFPPASKDVETVYCGELGKMG 336  
 Oy 53 SPEPTCOYIQCEPLSAPDLGIMNSHPLASF-----FTSACFTICSGTEGL 99

Db 337 IP-PACKATCO-----PLAQFSGNDRLEFGDGTGYGVSFRDCHSGYRR 381  
 Oy 100 IGKKTICSSSGIWSNPPICOKL 123  
 Db 382 EGVESLSLCKSDGTWSSKQPNCTKV 405

## RESULT 33

S57953  
 C4BP protein alpha chain precursor - rat  
 C:Species: Rattus norvegicus (Norway rat)  
 C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 20-Aug-1999  
 C:Accession: S57953  
 R:Hillary, A.; Thern, A.; Dahlback, B.  
 submitted to the EMBL Data Library, July 1995  
 A:Description: Molecular cloning of rat C4b-binding protein alpha- and beta-chains:  
 A:Reference number: S57953  
 A:Accession: S57953  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-558 <HIL>  
 A:Cross-references: EMBL:Z50051; NID:9899379; PTDN:CAA90391.1; PID:9899380  
 C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology  
 F:15-72/Domain: complement factor H repeat homology <FH1>  
 F:77-134/Domain: complement factor H repeat homology <FH2>  
 F:139-199/Domain: complement factor H repeat homology <FH3>  
 F:204-258/Domain: complement factor H repeat homology <FH4>  
 F:263-324/Domain: complement factor H repeat homology <FH5>  
 F:328-386/Domain: complement factor H repeat homology <FH6>  
 F:390-443/Domain: complement factor H repeat homology <FH7>  
 F:447-501/Domain: complement factor H repeat homology <FH8>

Query Match 19.9%; Score 146; DB 2; Length 558;  
 Best Local Similarity 28.6%; Pred. No. 1.3e-05;  
 Matches 42; Conservative 20; Mismatches 57; Indels 28; Gaps 7;

Oy 1 CEPLAPELLG-----TMDCTHFG-----NFSFSCQCAFSCSEGNITGIEETNC 45  
 Db 60 CKPLGKQWINTINACVKKSCNPGDLONGKVEYKTDPLFGSQIEFSGEYILGSSSTVCE 119  
 Oy 46 --GPFGNWSPEPTCOYIQCEPLSAPDLGIMNSH---PLASFSTTSACTFICSGTELI 100  
 Db 120 IOGKGVSMDSPLPECVIAKCG--MPD--ISGKINGRREFFYRSSVTVYACDDFTLL 175  
 Oy 101 GKKXTIC---ESSGIWNSPPICOKL 123  
 Db 176 GNASITCFVYNKTVGVWSPPPTCERI 202

## RESULT 34

I46001  
 C4b-binding protein alpha chain - bovine  
 C:Species: Bos primigenius taurus (cattle)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: I46001; S43190  
 R:Hillary, A.; Thern, A.; Dahlback, B.  
 J. Immunol. 153, 4190-4199, 1994  
 A:Title: Bovine C4b binding protein. Molecular cloning of the alpha- and beta-chains  
 A:Reference number: I46001; MUID:95015909  
 A:Accession: I46001  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-610 <HIL>  
 A:Cross-references: EMBL:Z31693; NID:9469117; PTDN:CAA83498.1; PID:9469118  
 C:Superfamily: C4b-binding protein alpha chain; complement factor H repeat homology  
 F:50-107/Domain: complement factor H repeat homology <FH1>  
 F:112-169/Domain: complement factor H repeat homology <FH2>  
 F:174-234/Domain: complement factor H repeat homology <FH3>  
 F:239-294/Domain: complement factor H repeat homology <FH4>  
 F:299-362/Domain: complement factor H repeat homology <FH5>  
 F:366-425/Domain: complement factor H repeat homology <FH6>



C:Superfamily: complement C3d/Epstein-Barr virus receptor; complement factor H repeat  
C:Keywords: alternative splicing; duplication; glycoprotein; transmembrane protein  
F:1-20/Domain: signal sequence #status predicted <SIG>  
F:21-1091/Product: complement receptor 2 (16-repeat form) #status predicted <NA11>  
F:21-658,718-1091/Product: complement receptor 2 (15-repeat form) #status predicted <  
F:23-82/Domain: complement factor H repeat homology <FH01>  
F:91-146/Domain: complement factor H repeat homology <FH02>  
F:154-210/Domain: complement factor H repeat homology <FH03>  
F:215-271/Domain: complement factor H repeat homology <FH04>  
F:275-342/Domain: complement factor H repeat homology <FH05>  
F:351-406/Domain: complement factor H repeat homology <FH06>  
F:410-465/Domain: complement factor H repeat homology <FH07>  
F:470-521/Domain: complement factor H repeat homology <FH08>  
F:528-592/Domain: complement factor H repeat homology <FH09>  
F:601-656/Domain: complement factor H repeat homology <FH10>  
F:660-716/Domain: complement factor H repeat homology <FH11>  
F:720-772/Domain: complement factor H repeat homology <FH12>  
F:777-837/Domain: complement factor H repeat homology <FH13>  
F:846-901/Domain: complement factor H repeat homology <FH14>  
F:909-965/Domain: complement factor H repeat homology <FH15>  
F:970-1026/Domain: complement factor H repeat homology <FH16>  
F:1034-1056/Domain: transmembrane #status predicted <TM>  
F:1057-1091/Domain: intracellular #status predicted <INT>  
F:121,127,294,372,622,698,858,881,919/Binding site: carbohydrate (Asn) (covalent) #st

Query Match 19.2%; Score 141; DB 1; Length 1091;  
Best Local Similarity 29.4%; Pred. No. 7,1e-05;  
Matches 35; Conservative 16; Mismatches 62; Indels 6; Gaps 3;

Oy 7 PELGTMDCTHPFGN--FSFSQCAFSCSEGTNLGIEETTCGPRGNMSPEPTGVITQC 63  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 911 PPPKPNNGNHGTGNIFARSPGSMILYSCDGYLLVGEALLLCTHCGTWSOPAPHCKEYNC 970

Oy 64 EPLSNPDL-GIMNCHPLASFTTSACFFICSGEHELIGKKRTICESSGINSNPPIQ 121  
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |  
Db 971 S--SPADDHGIDQKGLEPKRMRYOGAVVTLECDYGWLGESPGSQCSOSHQNMPPLAVCR 1027

RESULT 39

BBHU  
complement factor B precursor [validated] - human  
N:Alternate names: C3 convertase; C3 proactivator; glycine-rich beta-glycoprotein; heparin-binding protein  
N:Contains: alternative-complement-pathway C3/C5 convertase (EC 3.4.21.47) Bb fragment  
C:Species: Homo sapiens (man)  
C>Date: 19-feb-1984 #sequence-revision 05-Aug-1994 #text-change 08-Dec-2000  
C:Accession: S34075; A44622; A00934; A13988; A19947; B19947; B25971; S14339; A44628;  
R:Mejia, J.E.; Janm, I.; de la Salle, H.; Hauptmann, G.  
submitted to the EMBL Data Library, March 1993  
A:Reference number: S34075  
A:Accession: S34075  
A:Molecule type: mRNA  
A:Residues: 1-764 <MEJ>  
A:Cross-references: EMBL:X72875; NID:9297568; PIDN:CAAS1389.1; PID:9297569  
R:Moody, D.E.; Markham, A.F.; Ricker, A.T.; Goldberg, G.; Colten, H.R.  
Proc. Natl. Acad. Sci. U.S.A. 79, 5661-5665, 1982  
A>Title: Isolation of cDNA clones for the human complement protein factor B, a class  
A:Reference number: A44622; MUID:83039428  
A:Accession: A44622  
A:Molecule type: mRNA  
A:Residues: 467-546;550-595;752-764 <MOO>  
A:Cross-references: GB:J00185; GB:J00186  
A>Note: the authors translated the codon TAC at 519 as Thr; the nucleic acid translation  
R:Mole, J.E.; Anderson, J.K.; Davison, E.A.; Woods, D.E.  
J. Biol. Chem. 259, 3407-3412, 1984  
A>Title: Complete primary structure for the zymogen of human complement factor B.  
A:Reference number: A20751; MUID:84161997  
A:Accession: A00934  
A:Molecule type: protein; mRNA  
A:Residues: 26-764 <MO>  
A:Cross-references: GB:K01566  
A>Note: nucleic acid translation differs from the sequence shown in having 300-Leu, 3

A>Note: glycosylation sites were determined  
 R:Christie, D.L.; Gagnon, J.  
 Biochem. J. 209, 61-70, 1983  
 A:Title: Amino acid sequence of the Bb fragment from complement factor B. Sequence of th  
 A:Reference number: A19188; MUID:83204002  
 A:Contents: the final paper in a series documenting the sequence, glycosylation site, an  
 A:Accession: A19188  
 A:Molecule type: protein  
 A:Residues: 260-296, 'T', 298-764 <CHR>  
 R:Campbell, R.D.; Porter, R.R.  
 Proc. Natl. Acad. Sci. U.S.A. 80, 4464-4468, 1983  
 A:Title: Molecular cloning and characterization of the gene coding for human complement  
 A:Reference number: A19947; MUID:83273641  
 A:Accession: A19947  
 A:Molecule type: DNA  
 A:Residues: 346-764 <CAM>  
 A:Cross-references: GB:J00125  
 A:Accession: B19947  
 A:Molecule type: mRNA  
 A:Residues: 339-509 <CAI>  
 A:Cross-references: GB:J00126; NID:9187723; PIDN:AAA36226.1; PID:9553536  
 R:Wu, L.; Morley, B.J.; Campbell, R.D.  
 Cell 48, 331-342, 1987  
 A:Title: Cell-specific expression of the human complement protein factor B gene: evidenc  
 A:Reference number: A25971; MUID:87102880  
 A:Accession: B25971  
 A:Molecule type: DNA  
 A:Residues: 1-99 <MUI>  
 A:Cross-references: GB:M15082; NID:9187699; PIDN:AAA59625.1; PID:9553534  
 R:Niemann, W.A.; Bhown, A.S.; Miller, E.J.  
 Biochem. J. 274, 473-480, 1991  
 A:Title: The principal site of glycation of human complement factor B.  
 A:Reference number: S14339; MUID:91174758  
 A:Accession: S14339  
 A:Molecule type: protein  
 A:Residues: 270-329 <NIE>  
 A>Note: binding site for carbohydrate to lysine under artificial conditions  
 R:Morley, B.J.; Campbell, R.D.  
 EMBO J. 3, 153-157, 1984  
 A:Title: Internal homologies of the Ba fragment from human complement component factor H  
 A:Reference number: A44628; MUID:84158524  
 A:Accession: A44628  
 A>Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 16-225, 'F', 227-259 <MOR>  
 R:Schwaeble, W.; Lutting, B.; Sokolowski, T.; Estaller, C.; Weiss, E.H.; Meyer zum Busche  
 Immunobiology 188, 221-232, 1993  
 A:Title: Human complement factor B: functional properties of a recombinant zymogen of th  
 A:Reference number: I54409; MUID:94041399  
 A:Accession: I54409  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-764 <RES>  
 A:Cross-references: GB:S67310; NID:9452937; PIDN:AAD13989.1; PID:94261689  
 R:Horlrich, T.; Kim, S.; Matsumoto, M.; Watanabe, I.; Fujita, S.; Volanakis, J.E.  
 Mol. Immunol. 30, 1587-1592, 1993  
 A:Title: Human complement factor B: cDNA cloning, nucleotide sequencing, phenotypic con  
 A:Reference number: I57824; MUID:94067177  
 A:Accession: I57824  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-31, 'Q', 33-764 <RE2>  
 A:Cross-references: GB:L15702; NID:9291921; PIDN:AAA16820.1; PID:9291922  
 C:Comment: 292-Cys has a free sulhydryl.  
 C:Genetics:  
 A:Gene: GDB:BF  
 A:Cross-references: GDB:119726; OMIM:138470  
 A:Map position: 6p21.3-6p21.3  
 A:Introns: 21/3; 99/3; 346/1; 390/1; 424/1; 470/1; 502/3; 542/1; 593/2; 619/1; 652/3; 69  
 A>Note: the list of introns may be incomplete  
 C:Complex: complement factor B initially forms an inactive complex with complement facto  
 ment factor C3b forming active C3/C5 convertase; Ba is released

C:Function:  
 A:Description: Bb is a serine proteinase; C3/C5 convertase cleaves complement C3 alph  
 a1  
 A:Pathway: complement alternate pathway  
 C:Superfamily: complement C2; complement factor H repeat homology; trypsin homology;  
 C:Keywords: acute phase; complement alternate pathway; duplication; glycoprotein; hyd  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-764/Product: complement factor B #status experimental <MAT>  
 F:26-259/Product: complement factor Ba fragment #status experimental <BAR>  
 F:37-98/Domain: complement factor H repeat homology <PH1>  
 F:103-158/Domain: complement factor H repeat homology <PH2>  
 F:165-218/Domain: complement factor H repeat homology <PH3>  
 F:260-764/Product: C3/C5 convertase Bb fragment #status experimental <BBF>  
 F:268-458/Domain: von Willebrand factor type A repeat homology <VFA>  
 F:482-752/Domain: trypsin homology #status atypical <TRY>  
 F:37-76,62-98,103-145,131-158,165-205,191-218,478-596,511-527,599-615,656-682,695-725  
 F:122,142,285,378/Binding site: carbohydrate (asn) (covalent) #status experimental  
 F:259-260/Cleavage site: Arg-Lys (complement factor D) #status experimental  
 F:526,576,699/Active site: His, Asp, Ser #status experimental

Query Match 19.2% Score 140.5; DB 1; Length 764;  
 Best Local Similarity 25.2%; Pred. No. 5.5e-05;  
 Matches 35; Conservative 14; Mismatches 55; Indels 35; Gaps 4;

Qy 1 CEPLAEPLGTMDCTHPEGN-----FSSQCAFSCSEGTNLGIEETTCGPGMW 51  
 Db 98 CRAIHCP-----PHDFENGGEVWPSPYNNVSDSEIFHCYDGLTGLSANRRCQVNGRW 151  
 Qy 52 SSPEPTQVQICEPLSPDGLGIMNCSP-----LASFPSACTFCSESTELIGK 102  
 Db 152 SGQTAIC-----DNGAGYCSNPGIPIGTRKYGSQYRLEDSVTHCSRGLTLRGS 200  
 Qy 103 KRTICSSGIMSNPSPIQ 121  
 Db 201 QRRTCQSGSKSGTSPSCQ 219

RESULT 40  
 WNVZSP  
 apolipoprotein H homolog precursor - vaccinia virus  
 N:Alternate names: 35K secretory protein; C3L protein; virokin  
 C:Species: vaccinia virus  
 A:Note: host Homo sapiens (man)  
 C:Date: 31-Dec-1989 #sequence\_revision 30-Jun-1990 #text\_change 22-Jun-1999  
 C:Accession: A31005; B42504  
 R:Kotwal, G.J.; Moss, B  
 Nature 335, 176-178, 1988  
 A:Title: Vaccinia virus encodes a secretory polypeptide structurally related to compl  
 A:Reference number: A31005; MUID:88318974  
 A:Accession: A31005  
 A:Molecule type: DNA  
 A:Residues: 1-263 <ROT>  
 A:Cross-references: GB:X13166; NID:960690; PIDN:CAA31564.1; PID:960691  
 A:Experimental source: strain WR  
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.  
 Virology 179, 517-563, 1990  
 A:Title: Appendix to "The complete DNA sequence of vaccinia virus".  
 A:Reference number: A42501  
 A:Accession: A42501  
 A:Molecule type: DNA  
 A:Residues: 1-263 <GOE>  
 A:Cross-references: GB:M35027; NID:9335317; PIDN:AAA47997.1; PID:9335345  
 A:Experimental source: strain Copenhagen  
 R:Goebel, S.J.; Johnson, G.P.; Perkins, M.E.; Davis, S.W.; Winslow, J.P.; Paoletti, E.  
 Virology 179, 247-266, 1990  
 A:Title: The complete DNA sequence of vaccinia virus.  
 A:Reference number: A42531; MUID:91021027  
 A:Contents: annotation; possible protein-coding frames  
 A:Note: neither amino acid nor nucleotide sequence is given  
 C:Superfamily: herpesvirus complement control protein; complement factor H repeat hom  
 C:Keywords: duplication; extracellular protein  
 F:1-19/Domain: signal sequence #status predicted <SIG>





**This Page Blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:22:09 : Search time 85.16 Seconds  
(without alignments)  
392.017 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328  
Perfect score: 733  
Sequence: 1 CEPLAPELLGTMDCTHPEFGN.....MSNPSPICQKLDKSFMSIKE 132

Scoring table:  
BIOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 869228 seqs, 252910224 residues

Total number of hits satisfying chosen parameters: 869228

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Pending Patents AA-New\*  
1: /cgn2\_6/ptodata/2/paa/US06\_NEW\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/paa/US07\_NEW\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/paa/US08\_NEW\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/paa/US09\_NEW\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/paa/US10\_NEW\_COMB.pep:\*  
7: /cgn2\_6/ptodata/2/paa/US60\_NEW\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	733	100.0	372	US-09-791-537-152667	Sequence 152667, Ap
2	727	99.2	184	US-10-212-054-1437	Sequence 1437, Ap
3	727	99.2	372	US-09-791-537-22816	Sequence 22816, A
4	727	99.2	372	US-09-791-537-42637	Sequence 42637, A
5	727	99.2	372	US-09-791-537-84593	Sequence 84593, A
6	727	99.2	385	US-09-791-537-51391	Sequence 51391, A
7	724	98.8	372	US-09-791-537-42659	Sequence 42659, A
8	719	98.1	363	US-09-791-537-58446	Sequence 58446, A
9	719	98.1	385	US-09-791-537-113060	Sequence 113060, A
10	713	97.3	372	US-09-791-537-69658	Sequence 69658, A
11	713	97.3	385	US-09-791-537-53844	Sequence 53844, A
12	690	94.1	372	US-09-791-537-42655	Sequence 42655, A
13	690	94.1	372	US-09-791-537-42658	Sequence 42658, A
14	674	92.0	341	US-10-211-364-1171	Sequence 1171, Ap
15	674	92.0	341	US-10-212-054-1128	Sequence 1128, Ap
16	674	92.0	341	US-10-212-778-1158	Sequence 1158, Ap
17	599	81.7	376	US-09-791-537-132144	Sequence 132144, A
18	537	73.3	370	US-09-791-537-50403	Sequence 50403, A
19	535	73.0	372	US-09-791-537-81233	Sequence 81233, A
20	528	72.0	372	US-09-791-537-6693	Sequence 6693, Ap
21	518	70.7	323	US-09-791-537-53485	Sequence 53485, A
22	518	70.7	360	US-09-791-537-60503	Sequence 60503, A
23	518	70.7	372	US-09-791-537-37750	Sequence 37750, A
24	318	43.4	646	US-09-791-537-18300	Sequence 18300, A
25	315	43.0	740	US-09-791-537-32024	Sequence 32024, A
26	315	43.0	830	US-09-791-537-22819	Sequence 22819, A

27	315	43.0	830	5	US-09-791-537-35618	Sequence 35618, A
28	313	42.7	768	5	US-09-791-537-86038	Sequence 86038, A
29	309	42.2	616	5	US-09-791-537-33025	Sequence 33025, A
30	307	41.9	740	5	US-09-791-537-32023	Sequence 32023, A
31	306	41.7	754	5	US-09-791-537-14651	Sequence 14651, A
32	305.5	41.7	612	5	US-09-791-537-73416	Sequence 73416, A
33	305.5	41.7	618	5	US-09-791-537-81892	Sequence 81892, A
34	305.5	41.7	619	5	US-09-791-537-37749	Sequence 37749, A
35	300	40.9	611	5	US-09-791-537-118878	Sequence 118878, A
36	300	40.9	768	5	US-09-791-537-20989	Sequence 20989, A
37	300	40.9	768	5	US-09-791-537-37753	Sequence 37753, A
38	300	40.9	769	5	US-09-791-537-50409	Sequence 50409, A
39	299.5	40.9	549	5	US-09-791-537-50406	Sequence 50406, A
40	296	40.4	610	1	PCR-US02-23913-357	Sequence 357, App
41	296	40.4	610	5	US-09-791-537-22813	Sequence 22813, A
42	296	40.4	610	5	US-09-791-537-121834	Sequence 121834, A
43	296	40.4	610	6	US-10-205-823-357	Sequence 357, App
44	291	39.7	649	5	US-09-791-537-93873	Sequence 93873, A
45	290	39.6	646	5	US-09-791-537-84829	Sequence 84829, A

## ALIGNMENTS

```
RESULT 1
US-09-791-537-152667
; Sequence 152667, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biocomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 152667
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: X is an unknown amino acid
US-09-791-537-152667

Query Match      100.0%  Score 733;  DB 5;  Length 372;
Best Local Similarity 100.0%  Pred. No. 2.5e-57;
Matches 132;  Conservative 0;  Mismatches 0;  Indels 0;  Gaps 0;

QY 1 CEPLAPELLGTMDCTHPEFGNFSFSSQACSCSGTTLTGIEPTGCPFGNMSPEPTCOV 60
DB 197 CEPLAPELLGTMDCTHPEFGNFSFSSQACSCSGTTLTGIEPTGCPFGNMSPEPTCOV 256
QY 61 ICEPLAPELLGTMDCTHPEFGNFSFSSQACSCSGTTLTGIEPTGCPFGNMSPEPTCOV 120
DB 257 ICEPLAPELLGTMDCTHPEFGNFSFSSQACSCSGTTLTGIEPTGCPFGNMSPEPTCOV 316
QY 121 OKLDSFSMIKE 132
DB 317 OKLDSFSMIKE 328

RESULT 2
US-10-212-054-1437
; Sequence 1437, Application US/10212054
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P0212054
```

```
; CURRENT APPLICATION NUMBER: US/10/212,054
; CURRENT FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2164
; Prior application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1437
; LENGTH: 184
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-212-054-1437
```

```
Query Match          99.2%; Score 727; DB 6; Length 184;
Best Local Similarity 99.2%; Pred. No. 4.1e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 60
   |||||||||||
Db 9 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 68
QY 61 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
   |||||||||||
Db 69 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 128
QY 121 QKLDKSFMSIKE 132
   |||||||||||
Db 129 QKLDKSFMSIKE 140
```

```
RESULT 3
US-09-791-537-22816
; Sequence 22816, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 22816
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-22816
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 60
   |||||||||||
Db 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 256
QY 61 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
   |||||||||||
Db 257 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 316
QY 121 QKLDKSFMSIKE 132
   |||||||||||
Db 317 QKLDKSFMSIKE 328
```

```
RESULT 4
US-09-791-537-42657
; Sequence 42657, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
```

```
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42657
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Pan troglodytes
US-09-791-537-42657
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 60
   |||||||||||
Db 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 256
QY 61 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
   |||||||||||
Db 257 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 316
QY 121 QKLDKSFMSIKE 132
   |||||||||||
Db 317 QKLDKSFMSIKE 328
```

```
RESULT 5
US-09-791-537-84593
; Sequence 84593, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Biomolx, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 84593
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-84593
```

```
Query Match          99.2%; Score 727; DB 5; Length 372;
Best Local Similarity 99.2%; Pred. No. 8.6e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
QY 1 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 60
   |||||||||||
Db 197 CEPLAPELLGMDCTHPGNGFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOY 256
QY 61 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 120
   |||||||||||
Db 257 IQCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKKTICSSGIMNSPSPIC 316
QY 121 QKLDKSFMSIKE 132
   |||||||||||
Db 317 QKLDKSFMSIKE 328
```

```
RESULT 6
US-09-791-537-51391
; Sequence 51391, Application US/09791537
```

```
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 51391
; LENGTH: 385
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-791-537-51391
```

```
Query Match          99.2%; Score 727; DB 5; Length 385;
Best Local Similarity 99.2%; Pred. No. 8.9e-57;
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 60
    |||||||
DB 210 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 269
OY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 120
    |||||||
DB 270 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 329
OY 121 OKLDKSF5MIKE 132
    |||||||
DB 330 OKLDKSF5MIKE 341
```

RESULT 7

US-09-791-537-42659

```
; Sequence 42659, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 42659
; LENGTH: 372
; TYPE: PRP
; ORGANISM: Pongo pygmaeus
US-09-791-537-42659
```

```
Query Match          98.8%; Score 724; DB 5; Length 372;
Best Local Similarity 98.5%; Pred. No. 1.6e-56;
Matches 130; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 60
    |||||||
DB 197 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 256
OY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 120
    |||||||
DB 257 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 316
OY 121 OKLDKSF5MIKE 132
    |||||||
DB 317 OKLDKSF5MIKE 328
```

RESULT 8

US-09-791-537-58446

```
; Sequence 58446, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 58446
; LENGTH: 363
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-791-537-58446
```

```
Query Match          98.1%; Score 719; DB 5; Length 363;
Best Local Similarity 98.5%; Pred. No. 4.3e-56;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 60
    |||||||
DB 210 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 269
OY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 120
    |||||||
DB 270 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 329
OY 121 OKLDKSF5MIKE 132
    |||||||
DB 330 OKLDKSF5MIKE 341
```

RESULT 9

US-09-791-537-113060

```
; Sequence 113060, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Danzer, Derek
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 113060
; LENGTH: 385
; TYPE: PRP
; ORGANISM: Homo sapiens
US-09-791-537-113060
```

```
Query Match          98.1%; Score 719; DB 5; Length 385;
Best Local Similarity 98.5%; Pred. No. 4.6e-56;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 60
    |||||||
DB 210 CEPLAPELLGMDCTHPFGNFSFSSQAFSCSEGTNLGTIEETTCGPGNWSSPEPTCOV 269
OY 61 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 120
    |||||||
DB 270 IOCEPLASAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKKTTICSSGIMNSP1C 329
OY 121 OKLDKSF5MIKE 132
    |||||||
```

Db 330 OKLDRSFSMIKE 341

RESULT 10  
US-09-791-537-69658

; Sequence 69658, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 69658

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-69658

Query Match 97.3%; Score 713; DB 5; Length 372;

Best Local Similarity 97.0%; Pred. No. 1.5e-55;

Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 60

Db 197 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 256

QY 61 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 120

Db 257 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 316

QY 121 OKLDRSFSMIKE 132

Db 317 OKLDRSFSMIKE 328

RESULT 11

US-09-791-537-53844

; Sequence 53844, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 53844

; LENGTH: 385

; TYPE: PRT

; ORGANISM: Homo sapiens

US-09-791-537-53844

Query Match 97.3%; Score 713; DB 5; Length 385;

Best Local Similarity 97.0%; Pred. No. 1.6e-55;

Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 60

Db 210 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 269

QY 61 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 120

Db 270 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 329

QY 121 OKLDRSFSMIKE 132  
Db 330 OKLDRSFSMIKE 341

RESULT 12  
US-09-791-537-42655

; Sequence 42655, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 42655

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Macaca mulatta

US-09-791-537-42655

Query Match 94.1%; Score 690; DB 5; Length 372;

Best Local Similarity 92.4%; Pred. No. 1.7e-53;

Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 60

Db 197 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 256

QY 61 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 120

Db 257 IQCEPLAPDLGIMNCSHPLASFSTACTFTICSEGTLLIGKKTICSSGIMNSPSPIC 316

QY 121 OKLDRSFSMIKE 132

Db 317 OKLDRSFSMIKE 328

RESULT 13

US-09-791-537-42658

; Sequence 42658, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Biomomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

; TITLE OF INVENTION: METHODS OF USE THEREOF

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 42658

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Papio hamadryas

US-09-791-537-42658

Query Match 94.1%; Score 690; DB 5; Length 372;

Best Local Similarity 92.4%; Pred. No. 1.7e-53;

Matches 122; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

QY 1 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 60

Db 197 CEPLAPDLGMDCTHPGNGFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOY 256



Db 279 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 338  
QY 121 Q 121  
Db 339 Q 339

RESULT 17  
US-09-791-537-132144

; Sequence 132144, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 132144

; LENGTH: 376

; TYPE: PRT

; ORGANISM: Oryctolagus cuniculus

US-09-791-537-132144

Query Match 81.7%; Score 599; DB 5; Length 376;  
Best Local Similarity 80.3%; Pred. No. 2,1e-45;

Matches 106; Conservative 8; Mismatches 18; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 60  
Db 197 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 256  
QY 61 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 120  
Db 257 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 316  
QY 121 OKLDKSFMIKE 132  
Db 317 OKLDKSFMIKE 328

RESULT 18

US-09-791-537-50403

; Sequence 50403, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 50403

; LENGTH: 370

; TYPE: PRT

; ORGANISM: Bos taurus

US-09-791-537-50403

Query Match 73.3%; Score 537; DB 5; Length 370;  
Best Local Similarity 68.9%; Pred. No. 6,7e-40;

Matches 91; Conservative 19; Mismatches 22; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 60  
Db 197 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 256

Db 197 CVPLAPKLGTMCTHPLGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 256

QY 61 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 120

Db 257 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 316

QY 121 OKLDKSFMIKE 132

Db 317 OKLDKSFMIKE 328

RESULT 19

US-09-791-537-81233

; Sequence 81233, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 81233

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Rattus norvegicus

US-09-791-537-81233

Query Match 73.0%; Score 535; DB 5; Length 372;  
Best Local Similarity 71.2%; Pred. No. 1e-39;

Matches 94; Conservative 15; Mismatches 23; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 60  
Db 197 CEPLAPELGTMDCTHFGNFSFSSQCAFSCSGTNTLGIEETTCGPGNWSPEPTCOV 256  
QY 61 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 120  
Db 257 IQCEPLSAPDLGIMNCSHPLASFSFTSACTFICSEGTLLIGKKTICSSGIMNSNPIC 316  
QY 121 OKLDKSFMIKE 132  
Db 317 OKLDKSFMIKE 328

RESULT 20

US-09-791-537-6693

; Sequence 6693, Application US/09791537

; GENERAL INFORMATION:

; APPLICANT: Blonomix, Inc.

; APPLICANT: Debe, Derek

; APPLICANT: Danzer, Joseph

; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME

; FILE REFERENCE: 261/210

; CURRENT APPLICATION NUMBER: US/09/791,537

; CURRENT FILING DATE: 2001-02-22

; NUMBER OF SEQ ID NOS: 153055

; SOFTWARE: PatentIn version 3.0

; SEQ ID NO 6693

; LENGTH: 372

; TYPE: PRT

; ORGANISM: Rattus sp

US-09-791-537-6693

Query Match 72.0%; Score 528; DB 5; Length 372;  
Best Local Similarity 70.5%; Pred. No. 4,2e-39;

Matches 93; Conservative 15; Mismatches 24; Indels 0; Gaps 0;





Best Local Similarity 42.1%; Pred. No. 3.5e-20;  
Matches 51; Conservative 22; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPFNMSSPEPTCOV 60
    |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 CQHLAEPSEGTMDCVHPLTAFAYGSSCKFECPQGYRVRGIDMLRCIDSGHMSAPLPTCEA 383
OY 61 IQCEPLSAPDLGIMNCSHPLASFSTTACTFTICSEGTTELIGKKRTICSSGIMNSPPTIC 120
    :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 384 LQCDLPAPKAOVNCSPHPCAFRYOSTCSFTCDGSSIVGASVLCLETGMSAPAPVC 443
OY 121 Q 121
    |
Db 444 Q 444
```

RESULT 25  
US-09-791-537-32024  
Sequence 32024, Application US/09791537

GENERAL INFORMATION:  
APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 32024

LENGTH: 740

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-32024

Query Match 43.0%; Score 315; DB 5; Length 740;  
Best Local Similarity 43.9%; Pred. No. 7.5e-20;  
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPFNMSSPEPTCOV 60
    |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 CQHLAEPSEGTMDCVHPLTAFAYGSSCKFECPQGYRVRGIDMLRCIDSGHMSAPLPTCEA 321
OY 61 IQCEPLSAPDLGIMNCSHPLASFSTTACTFTICSEGTTELIGKKRTICSSGIMNSPPTIC 120
    :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 322 ISCEPLSPVHGSMDCSPSLRAFQYDYNCSFRCAGFMRLGADIVRCDNLQGTAPAPVC 381
OY 121 QKL 123
    |
Db 382 QAL 384
```

RESULT 26  
US-09-791-537-22819  
Sequence 22819, Application US/09791537

GENERAL INFORMATION:  
APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 22819

LENGTH: 830

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-22819

Query Match 43.0%; Score 315; DB 5; Length 830;  
Best Local Similarity 43.9%; Pred. No. 8.5e-20;  
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPFNMSSPEPTCOV 60
    |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 CQHLAEPSEGTMDCVHPLTAFAYGSSCKFECPQGYRVRGIDMLRCIDSGHMSAPLPTCEA 383
OY 61 IQCEPLSAPDLGIMNCSHPLASFSTTACTFTICSEGTTELIGKKRTICSSGIMNSPPTIC 120
    :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 384 LQCDLPAPKAOVNCSPHPCAFRYOSTCSFTCDGSSIVGASVLCLETGMSAPAPVC 443
OY 121 QKL 123
    |
Db 444 QAL 446
```

RESULT 27  
US-09-791-537-35618  
Sequence 35618, Application US/09791537

GENERAL INFORMATION:  
APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 35618

LENGTH: 830

TYPE: PRT

ORGANISM: Homo sapiens

US-09-791-537-35618

Query Match 43.0%; Score 315; DB 5; Length 830;  
Best Local Similarity 43.9%; Pred. No. 8.5e-20;  
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAPELLGTMDCTHPPGNFSSQCAFSCSEGTNLGIEETTCGPFNMSSPEPTCOV 60
    |||||:| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 CQHLAEPSEGTMDCVHPLTAFAYGSSCKFECPQGYRVRGIDMLRCIDSGHMSAPLPTCEA 383
OY 61 IQCEPLSAPDLGIMNCSHPLASFSTTACTFTICSEGTTELIGKKRTICSSGIMNSPPTIC 120
    :|||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:| |||:|
Db 384 LQCDLPAPKAOVNCSPSLRAFQYDYNCSFRCAGFMRLGADIVRCDNLQGTAPAPVC 443
OY 121 QKL 123
    |
Db 444 QAL 446
```

RESULT 28  
US-09-791-537-86038  
Sequence 86038, Application US/09791537

GENERAL INFORMATION:  
APPLICANT: Biomimix, Inc.

APPLICANT: Debe, Derek

APPLICANT: Danzer, Joseph

TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS

FILE REFERENCE: 261/210

CURRENT APPLICATION NUMBER: US/09/791,537

CURRENT FILING DATE: 2001-02-22

NUMBER OF SEQ ID NOS: 153055

SOFTWARE: PatentIn version 3.0

SEQ ID NO 86038

LENGTH: 768

```
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-09-791-537-86038
```

```
Query Match          42.7%; Score 313; DB 5; Length 766;
Best Local Similarity 43.9%; Pred. No. 1.2e-19;
Matches 54; Conservative 17; Mismatches 52; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPGNSFSSQAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 262 CGSELPPLGHTMDCTHPLAFAIDSSCKECPGYRMRSDDIHCYDSCGWSPELPTCEA 321
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKTKICSSGIMNSPITC 120
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 322 IACEPLSLHGSMDCFPSTGANGYNSCTCFRCTESFVLMGNDIHCADLGOWTAPAPVC 381
OY 121 QKL 123
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 382 EAL 384
```

```
RESULT 29
US-09-791-537-32025
```

```
; Sequence 32025, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32025
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-32025
```

```
Query Match          42.2%; Score 309; DB 5; Length 616;
Best Local Similarity 41.5%; Pred. No. 2.1e-19;
Matches 51; Conservative 23; Mismatches 49; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPGNSFSSQAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 200 CGEELPQHVLMNCSPHLGNFSFNSQCSFHCYDGVNPGSKLECLASGIMTNKPPQCLA 259
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKTKICSSGIMNSPITC 120
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 260 ACCPPLKIPERGMMTCLHSAKAFOHSSCSFCEGFAVGPVEVQCTASGVWTAAPAVC 319
OY 121 QKL 123
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 320 KAL 322
```

```
RESULT 30
US-09-791-537-32023
```

```
; Sequence 32023, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
```

```
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 32023
; LENGTH: 740
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-791-537-32023
```

```
Query Match          41.9%; Score 307; DB 5; Length 740;
Best Local Similarity 40.7%; Pred. No. 3.9e-19;
Matches 50; Conservative 24; Mismatches 49; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPGNSFSSQAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 200 CGEELPQHVLMNCSPHLGNFSFNSQCSFHCYDGVNPGSKLECLASGIMTNKPPQCLA 259
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKTKICSSGIMNSPITC 120
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 260 ACCPPLKIPERGMMTCLHSAKAFOHSSCSFCEGFAVGPVEVQCTASGVWTAAPAVC 319
OY 121 QKL 123
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 320 KAL 322
```

```
RESULT 31
```

```
US-09-791-537-14651
; Sequence 14651, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 14651
; LENGTH: 754
; TYPE: PRT
; ORGANISM: Canis familiaris
US-09-791-537-14651
```

```
Query Match          41.7%; Score 306; DB 5; Length 754;
Best Local Similarity 40.5%; Pred. No. 4.8e-19;
Matches 49; Conservative 24; Mismatches 48; Indels 0; Gaps 0;
```

```
OY 1 CEPLAPELLGMDCTHPGNSFSSQAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 185 CGDFPLPQHVLMNCSPHLGNFSFNSQCSFHCYDGVNPGSKLECLASGIMTNKPPQCLA 244
OY 61 IOCEPLSAPDLGIMNCSHPLASFSTSACTFICSEGTTELIGKKTKICSSGIMNSPITC 120
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 245 IOCPPLKIPERGMMTCLHSAKAFOHSSCSFCEGFAVGPVEVQCTASGVWTAAPAVC 304
OY 121 Q 121
    | : | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 305 E 305
```

```
RESULT 32
```

```
US-09-791-537-73416
; Sequence 73416, Application US/09791537
; GENERAL INFORMATION:
```

```
; APPLICANT: Biomomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY ME
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
```



```

: APPLICANT: Bionomix, Inc.
: APPLICANT: Debeo, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 20989
: LENGTH: 768
: TYPE: PRT
: ORGANISM: Mus musculus
: US-09-791-537-20989

```

Query Match	40.9%;	Score 300;	DB 5;	Length 768;
Best Local Similarity	42.3%;	Pred. No. 1.7e-18;		
Matches 52;	Conservative 17;	Mismatches 54;	Indels 0;	Gaps 0;

QY	1	C E P L E A P E L G M O C T H P P G N F S F S C O A F S C S E G T N L G I E E T T C G F F G M W S S P E R T C Q Y	60
		:  :   :   :   :   :                 :   :	
D b	200	G K A N N I P D H V L M N C S H P L G E F S F S Q C T F S C A E E Y E I L D P G E L O C L A S G I W T N N P C D A	259
QY	61	I Q C E P L A P D L G I M N C S H P L A S F S T S A C T F I C S E G T E L L G K K T T C E S S G I W S N S P I C	120
		:  :   :   :   :   :   :   :   :   :   :   :	
D b	260	V O C S L E A P P H G T M A C M H P I A F A Y D S S C K F E C P O G Y R A R G S N T L H C T G S G W S E P L P T C	319
QY	121	O K L	123
	:	:	:
D b	320	E A I	322

```

RESULT 37
US-09-791-537-37753
: Sequence 37753, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: TITLE OF INVENTION: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMB
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 37753
: LENGTH: 768
: TYPE: PRT
: ORGANISM: Mus musculus
: US-09-791-537-37753

```

Query Match	40.9%	Score 300;	DB 5;	Length 768;
Best Local Similarity	42.3%	Pred. No. 1.7e-18;		
Matches	52;	Conservative	17;	Mismatches 54;
			Indels	0;
			Gaps	0;

[illegible]

```

US-09-791-537-50409
: Sequence 50409, Application US/09791A537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 50409
: LENGTH: 769
: TYPE: PRT
: ORGANISM: Ovis aries
US-09-791-537-50409

```

Query Match	40.9%	Score 300;	DB 5;	Length 769;
Best Local Similarity	39.8%;	Pred. No. 1.7e-18;		
Matches 49;	Conservative 24;	Mismatches 50;	Indels 0;	Gaps 0;

Qy	1	CEPLAPLGLTDCDTHPPGNSFFSSQCAFSCSEETNLTLGIEETCGFGNMSPEPPQV	60
		:  :	
Db	200	CGEEDLPQNMHNCNHLPGNFFSQCSFRHBEYALNGRELECLASGLTMTNSPQCA	259
Qy	61	IQCEPLSPADGLIMNCNHLPLASFSFTSACFTICEGIELLGKKKTTICESSGINSPPIC	120
		:  :	
	260	VQCPALKSPGGNSCSFSAKAFHOSSCSPSCGEGTLVGPVHCHTALGAVTAPVPC	319

```

RESULT 39
US-09-791-537-50406
: Sequence 50406, Application US/09791537
: GENERAL INFORMATION:
: APPLICANT: Bionomix, Inc.
: APPLICANT: Debe, Derek
: APPLICANT: Danzer, Joseph
: TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBERS
: TITLE OF INVENTION: METHODS OF USE THEREOF
: FILE REFERENCE: 261/210
: CURRENT APPLICATION NUMBER: US/09/791,537
: CURRENT FILING DATE: 2001-02-22
: NUMBER OF SEQ ID NOS: 153055
: SOFTWARE: patentIn version 3.0
: SEQ ID NO 50406
: LENGTH: 549
: TYPE: PRT
: ORGANISM: Rattus norvegicus
: US-09-791-537-50406

```

Query Match	40.98;	Score 299.5;	DB 5;	Length 549;
Best Local Similarity	41.08;	Pred. No. 1.3e-18;		
Matches 50;	Conservative 23;	Mismatches 48;	Indels 1;	Gaps 1;

**Qy**    1    CEPLEAEDELGDTHTFGNFSFSSCOAFSCSEETNLTLGEETT - CEFPMNSMSEPTECQ    59  
          | : : : : : | : : : : : | : : : : : | : : : : : |  
**Dd**    180   QCEDEYDPDGLSINCTHFFGFLEFSTYSSCSFSECRGYVPSMETVTRCTISGEWSPAPACH    239  
          | : : : : : | : : : : : | : : : : : | : : : : : |  
**Qy**    60   VIOCEPRLAPDLGMNCSHPLASFSPFSACTFLTCSEGTLLGKKKTKTCESSGIIMSNPSPI    119  
          | : : : : : | : : : : : | : : : : : | : : : : : |  
**Dd**    240   VVEKALTOPAHGVKRCCSNPGSTPMWNTCTCFDCBBEYRRRVGAQNLOCTYSGWDMKEPS    299  
          | : : : : : | : : : : : | : : : : : | : : : : : |  
**Qy**    120   CQ    121  
          | :  
**Dd**    300   CK    301



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 7, 2002, 10:20:38 : Search time 304.64 seconds  
(without alignments)  
152.513 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328

Perfect score: 733  
Sequence: 1 CEPLAEAEELGTMDCTHFGN.....WSNPSPICQKDKSFSMIKE 132

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 3502263 seqs, 351980561 residues

Total number of hits satisfying chosen parameters: 3502263

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Pending\_Patents\_AA\_Main:\*

1: /cgn2\_6/ptodata/2/paa/PCrUS\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/paa/US06\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/paa/US07\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/paa/US08\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/paa/US081\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/paa/US082\_COMB.pep.\*  
7: /cgn2\_6/ptodata/2/paa/US083\_COMB.pep.\*  
8: /cgn2\_6/ptodata/2/paa/US084\_COMB.pep.\*  
9: /cgn2\_6/ptodata/2/paa/US085\_COMB.pep.\*  
10: /cgn2\_6/ptodata/2/paa/US086\_COMB.pep.\*  
11: /cgn2\_6/ptodata/2/paa/US087\_COMB.pep.\*  
12: /cgn2\_6/ptodata/2/paa/US088\_COMB.pep.\*  
13: /cgn2\_6/ptodata/2/paa/US089\_COMB.pep.\*  
14: /cgn2\_6/ptodata/2/paa/US090\_COMB.pep.\*  
15: /cgn2\_6/ptodata/2/paa/US091\_COMB.pep.\*  
16: /cgn2\_6/ptodata/2/paa/US092\_COMB.pep.\*  
17: /cgn2\_6/ptodata/2/paa/US093\_COMB.pep.\*  
18: /cgn2\_6/ptodata/2/paa/US094\_COMB.pep.\*  
19: /cgn2\_6/ptodata/2/paa/US095\_COMB.pep.\*  
20: /cgn2\_6/ptodata/2/paa/US097\_COMB.pep.\*  
21: /cgn2\_6/ptodata/2/paa/US098\_COMB.pep.\*  
22: /cgn2\_6/ptodata/2/paa/US099\_COMB.pep.\*  
23: /cgn2\_6/ptodata/2/paa/US100\_COMB.pep.\*  
24: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*  
25: /cgn2\_6/ptodata/2/paa/US101\_COMB.pep.\*  
26: /cgn2\_6/ptodata/2/paa/US60\_COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	733	100.0	372	15	US-09-119-209-2
2	727	99.2	184	21	US-09-760-443-1437
3	727	99.2	184	21	US-09-760-475-3252
4	727	99.2	372	1	PCT-US01-26675-3
5	727	99.2	382	21	US-09-760-475-2123
6	727	99.2	1078	26	US-60-207-315-428
7	727	99.2	1078	26	US-60-212-659-523

Result No.	Score	Query Match	Length	ID	Description
8	727	99.2	1078	26	US-60-230-435-1751
9	713	97.3	385	1	PCT-US92-03970-2
10	713	97.3	385	1	PCT-US94-00909-2
11	713	97.3	385	4	US-08-008-459-2
12	713	97.3	385	7	US-08-340-539-2
13	713	97.3	385	8	US-08-410-569-2
14	674	92.0	341	21	US-09-758-449-1158
15	674	92.0	341	21	US-09-760-443-1328
16	518	70.7	372	15	US-09-119-209-4
17	315	43.0	830	1	PCT-US94-09395-4
18	315	43.0	830	8	US-08-449-687B-4
19	315	43.0	830	24	US-10-020-141-10
20	306	41.7	700	26	US-60-207-315-467
21	296	40.4	610	1	PCT-US99-28965-19
22	296	40.4	610	10	US-08-657-753-2
23	296	40.4	610	11	US-08-770-433-3
24	296	40.4	610	16	US-09-266-091A-2
25	296	40.4	610	16	US-09-266-091A-2
26	296	40.4	610	21	US-09-784-356-122
27	296	40.4	610	22	US-09-802-640-36
28	296	40.4	610	22	US-09-857-670-19
29	296	40.4	610	24	US-10-021-660-122
30	268	36.6	355	1	PCT-US01-01239-1145
31	268	36.6	355	21	US-09-764-902-1145
32	255	34.8	309	1	PCT-US01-01332-615
33	255	34.8	309	21	US-09-764-875-615
34	243	33.2	135	26	US-60-196-718-4486
35	238	32.5	208	21	US-09-760-475-2126
36	236	32.2	207	21	US-09-760-443-1321
37	236	32.2	207	21	US-09-760-475-3255
38	236	32.2	207	21	US-09-760-498-920
39	222	30.3	138	1	PCT-US01-01239-1688
40	222	30.3	138	1	PCT-US01-01332-938
41	222	30.3	138	21	US-09-764-875-938
42	222	30.3	138	21	US-09-764-902-1688
43	219.5	29.1	175	1	PCT-US01-03800A-1802
44	199	27.9	107	26	US-60-196-718-5358
45	198	27.0	106	26	US-60-196-710-5336

#### ALIGNMENTS

RESULT 1

US-09-119-209-2

Sequence 2, Application US/09119209

GENERAL INFORMATION:

APPLICANT: LASKY, LAURENCE A.

APPLICANT: STACHELL, SCOTT E.

APPLICANT: ROSEN, STEVEN D.

APPLICANT: SINGER, MARK S.

APPLICANT: YEDNICK, TED A.

TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 1 DNA Way

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: WinPatIn (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/119, 209

FILING DATE: 20-JUL-1998

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/513278

FILING DATE: 10-AUG-1995

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/059027
?
? FILING DATE: 6-MAY-1993
?
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/786149
? FILING DATE: 31-OCT-1991
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 07/315015
? FILING DATE: 23-FEB-1989
?
? ATTORNEY/AGENT INFORMATION:
?
? NAME: Love, Richard B.
? REGISTRATION NUMBER: 34,659
? REFERENCE/DOCKET NUMBER: P0565D1C3
?
? TELECOMMUNICATION INFORMATION:
?
? TELEPHONE: 650/225-5530
?
? TELEFAX: 650/952-9881
?
? INFORMATION FOR SEQ ID NO: 2:
?
? SEQUENCE CHARACTERISTICS:
?
? LENGTH: 372 amino acids
?
? TYPE: Amino Acid
?
? TOPOLOGY: Linear
?
US-09-119-209-2

```

Query Match	100.0%	Score 733:	DB 15:	Length 372:
Best Local Similarity	100.0%	Pred. NO. 6.6e-63:		
Matches 132:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	1	CEPLEAPELGLMDCHHPGNGFSSSQCAFSCSEGNLNGIEETTCGPFQNMSSPEPTCOV	60	
Db	197	CEPLEAPELGLMDCHHPGNGFSSSQCAFSCSEGNLNGIEETTCGPFQNMSSPEPTCOV	256	
QY	61	IQCEPLAPDGLIMNCSHPLASFSTSACTFICSEGTETLIGKKRTICSSSGIWSNPSPIC	120	
Db	257	IQCEPLAPDGLIMNCSHPLASFSTSACTFICSEGTETLIGKKRTICSSSGIWSNPSPIC	316	
QY	121	QKLDKSFMSIKE	132	
Db	317	QKLDKSFMSIKE	328	

```

RESULT      2
US-09-760-443-1437
: Sequence 1437, Application US/09760443
:
: GENERAL INFORMATION:
: APPLICANT: Rosen et al.
: TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
: FILE REFERENCE: PT12
:
: CURRENT APPLICATION NUMBER: US/09/760,443
: CURRENT FILING DATE: 2001-01-16
: Prior application data removed - refer to PALM or file wrapper
: NUMBER OF SEQ ID NOS: 2164
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 1437
:
: LENGTH: 184
:
: TYPE: PRT
:
: ORGANISM: Homo sapiens
: US-09-760-443-1437

```

Query Match	99.28%	Score 727:	DB 21:	Length 184:
Best Local Similarity	99.28%	Pred. No. 1.2e-62:		
Matches 131: Conservative	0:	Mismatches 1:	Indels 0:	Gaps 0:
QY	1	CEPLEAPELGMDCTHPFGNFSFSSOCASFSCSEGNLTGIEETTCGPGNMSSPEPTQY	60	
Db	9	CEPLEAPELGMDCTHPGNFSFSSQCAFSSCBSTNLGIEETTCGPGNMSSPEPTQY	68	
QY	61	IOCEPLTAPDGIINNCCHPLASFSTTACATPICESEGTFLIKRKTICSSGIWNSPDTIC	120	
Db	69	IOCEPLTAPDGIINNCCHPLASFSTTACATPICESEGTFLIKRKTICSSGIWNSPDTIC	128	
QY	121	QKLDKSFMSIKE	132	

```

Db          129  GRKDRSEFSMIKE  140
|||||
RESULT      3
US-09-760-475-3252
; Sequence 3252, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PAM or file wrapper
; NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3252
; LENGTH: 184
; TYPE: prt
; ORGANISM: Homo sapiens
US-09-760-475-3252

```

Query Match	99.2%	Score 727	DB 21	Length 184
Best Local Similarity	99.2%	Pred. No. 1.2e-62		
Matches 131: Conservative	0	Mismatches 1	Indels 0	Gaps 0
QY	1	CEPLEAPELGTMDCHNPGNFSSQCAFSCSEGNLNGIEETTCGPGNMSSPEPTCOV	60	
Db	9	CEPLEAPELGTMDCHNPLGNFSSQCAFSCSEGNLNGIEETTCGPGNMSSPEPTCOV	68	
QY	61	IQCEPLSAPADLGIMNCSHPLASFSTSACTFICSEGTETLIGKKKTIICSSSGIWSNPSPIC	120	
Db	69	IQCEPLSAPADLGIMNCSHPLASFSTSACTFICSEGTETLIGKKKTIICSSSGIWSNPSPIC	128	
QY	121	OKLDKSFESMIKE	132	
Db	129	OKLDKSFESMIKE	140	

```

RESULT      4
PCT-US01-26675-3
; Sequence 3, Application PC/TUS0126675
; GENERAL INFORMATION:
; APPLICANT: Genaisance Pharmaceuticals, Inc.
; APPLICANT: Anastasio, Alison E
; APPLICANT: Bieglecki, Karyn M
; APPLICANT: Kliehm, Stefanie E
; APPLICANT: Koshy, Beena
; APPLICANT: Kumar, Anant Madan
; TITLE OF INVENTION: HAPLOTYPES OF THE SELL GENE
; FILE REFERENCE: SELL MMH116-PCT
; CURRENT APPLICATION NUMBER: PCT/US01/26675
; PRIORITY FILING DATE: 2001-08-27
; PRIOR APPLICATION NUMBER: 60/228,262
; PRIORITY FILING DATE: 2000-08-25
; NUMBER OF SEQ ID NOS: 101
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
; PCT-US01-26675-3

```

Query Match	99.28%	Score 727	DB 1	Length 372
Best Local Similarity	99.28%	Pred. No.	2.5e-67	
Matches 131	Conservative	0	Mismatches 1	Indels 0
Gaps				0
QY	1	CEPLEAEPLGMDCTHFGNFSSSCAFCSCSECTNLGTLEETTCGPGNMSSDEPTCOV	60	
197	CEPLEAEPLGMDCTHFGNFSSSCAFCSCSECTNLGTLEETTCGPGNMSSDEPTCOV	256		



QY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 120  
|||||  
Db 257 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 316  
QY 121 OKLDKSFMSIKE 132  
|||||  
Db 317 OKLDKSFMSIKE 328

## RESULT 5

US-09-760-475-2123  
; Sequence 2123, Application US/09760475  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PT249  
; CURRENT APPLICATION NUMBER: US/09/760,475  
; CURRENT FILING DATE: 2001-01-16  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 4122  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2123  
; LENGTH: 382  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-760-475-2123

Query Match 99.2%; Score 727; DB 21; Length 382;  
Best Local Similarity 99.2%; Pred. No. 2,1e-62;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 60  
|||||  
Db 207 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 266  
QY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 120  
|||||  
Db 267 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 326  
QY 121 OKLDKSFMSIKE 132  
|||||  
Db 327 OKLDKSFMSIKE 338

## RESULT 6

US-60-207-315-428  
; Sequence 428, Application US/60207315  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: C1000601  
; CURRENT APPLICATION NUMBER: US/60/207,315  
; CURRENT FILING DATE: 2000-05-30  
; NUMBER OF SEQ ID NOS: 528  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 428  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-207-315-428

Query Match 99.2%; Score 727; DB 26; Length 1078;  
Best Local Similarity 99.2%; Pred. No. 8,1e-62;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 60  
|||||  
Db 779 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 838

QY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 120  
|||||  
Db 839 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 898  
QY 121 OKLDKSFMSIKE 132  
|||||  
Db 899 OKLDKSFMSIKE 910

## RESULT 7

US-60-212-659-523  
; Sequence 523, Application US/60212659  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: C1000674  
; CURRENT APPLICATION NUMBER: US/60/212,659  
; CURRENT FILING DATE: 2000-06-19  
; NUMBER OF SEQ ID NOS: 879  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 523  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-212-659-523

Query Match 99.2%; Score 727; DB 26; Length 1078;  
Best Local Similarity 99.2%; Pred. No. 8,1e-62;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 60  
|||||  
Db 779 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 838  
QY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 120  
|||||  
Db 839 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 898  
QY 121 OKLDKSFMSIKE 132  
|||||  
Db 899 OKLDKSFMSIKE 910

## RESULT 8

US-60-230-435-1751  
; Sequence 1751, Application US/60230435  
; GENERAL INFORMATION:  
; APPLICANT: Beasley, Ellen  
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS,  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND  
; FILE REFERENCE: C1000768  
; CURRENT APPLICATION NUMBER: US/60/230,435  
; CURRENT FILING DATE: 2000-09-06  
; NUMBER OF SEQ ID NOS: 2991  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1751  
; LENGTH: 1078  
; TYPE: PRT  
; ORGANISM: HUMAN  
US-60-230-435-1751

Query Match 99.2%; Score 727; DB 26; Length 1078;  
Best Local Similarity 99.2%; Pred. No. 8,1e-62;  
Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 60  
|||||  
Db 779 CEPLAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKKTTICSSGIMNSPSPIC 838

QY 61 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 120  
DB 839 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 898  
QY 121 QKDKSFSMIKE 132  
DB 899 QKDKSFSMIKE 910

## RESULT 9

PCT-US92-03970-2  
Sequence 2, Application PC/TUS9203970  
GENERAL INFORMATION:  
APPLICANT: Dana-Farber Cancer Institute, Inc.  
TITLE OF INVENTION: LEUKOCYTE-ASSOCIATED CELL SURFACE  
PROTEIN  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: U.S.A.  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/03970  
FILING DATE: 19920513  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Helne, Holliday C.  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: DCT-152Bq9  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
TELEX: 940675  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US92-03970-2

Query Match 97.3%; Score 713; DB 1; Length 385;  
Best Local Similarity 97.0%; Pred. No. 6.1e-61;  
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 60  
DB 210 CEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 269  
QY 61 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 120  
DB 270 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 329  
QY 121 QKDKSFSMIKE 132  
DB 330 QKDKSFSMIKE 341

RESULT 10  
PCT-US94-00909-2  
Sequence 2, Application PC/TUS9400909  
GENERAL INFORMATION:  
APPLICANT:

TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS BLOCKING  
AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US94/00909  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/983,606  
FILING DATE: 30-NOV-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/962,483  
FILING DATE: 02-APR-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,608  
FILING DATE: 03-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/737,092  
FILING DATE: 29-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/730,503  
FILING DATE: 08-JUL-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/700,773  
FILING DATE: 15-MAY-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/313,109  
FILING DATE: 21-FEB-1989  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US94-00909-2

Query Match 97.3%; Score 713; DB 1; Length 385;  
Best Local Similarity 97.0%; Pred. No. 6.1e-61;  
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 60  
DB 210 CEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 269  
QY 61 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 120  
DB 270 IOCEPLAPDLGIMNCNCHPLASFTSACTFICSEGTETLIGKKTICSSGIMNSPSPIC 329  
QY 121 QKDKSFSMIKE 132  
DB 330 QKDKSFSMIKE 341

RESULT 11  
US-08-008-459-2  
Sequence 2, Application US/08008459  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
CITY: Boston

```
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/008,459
FILING DATE: 25-JAN-1993
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/737,092
FILING DATE: 29-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-008-459-2

Query Match          97.3%; Score 713; DB 4; Length 385;
Best Local Similarity 97.0%; Pred. No. 6,1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

```
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSER: Weingarten, Schurgin, Gagnebin & Hayes
STREET: Ten Post Office Square
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/340,539
FILING DATE: 16-NOV-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/008,459
FILING DATE: 25-JAN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/983,606
FILING DATE: 30-NOV-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/862,483
FILING DATE: 02-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/770,608
FILING DATE: 03-OCT-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/730,503
FILING DATE: 08-JUL-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/700,773
FILING DATE: 15-MAY-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/313,109
FILING DATE: 21-FEB-1989
ATTORNEY/AGENT INFORMATION:
NAME: Heine, Holliday C.
REGISTRATION NUMBER: 34,346
REFERENCE/DOCKET NUMBER: DFCI-318XX
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-2290
TELEFAX: (617) 451-0313
TELEX: 940675
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-340-539-2

Query Match          97.3%; Score 713; DB 7; Length 385;
Best Local Similarity 97.0%; Pred. No. 6,1e-61;
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```

Db 330 QKDKSFSMIKE 341

|||||

RESULT 13  
US-08-410-569-2  
Sequence 2 Application US/08410569  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Spertini, Olivier G.  
TITLE OF INVENTION: LEUKOCYTE ADHESION MOLECULE-1 (LAM-1)  
TITLE OF INVENTION: AND LIGAND THEREOF  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/410,569  
FILING DATE:  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/770,608  
FILING DATE: 03-OCT-1991  
APPLICATION NUMBER: US 07/700,773  
FILING DATE: 15-MAY-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Heine, Holliday C.  
REGISTRATION NUMBER: 34,346  
REFERENCE/DOCKET NUMBER: DFCG-152EX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
TELEX: 940675

INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-410-569-2

Query Match 97.3%; Score 713; DB 8; Length 385;  
Best Local Similarity 97.0%; Pred. No. 6.1e-61;  
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60  
Db 210 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 269

QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 120  
Db 270 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 329

QY 121 QKDKSFSMIKE 132  
Db 330 QKDKSFSMIKE 341

RESULT 14  
US-09-758-449-1158  
Sequence 1158, Application US/09758449  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.

;; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
;; FILE REFERENCE: PM026  
;; CURRENT APPLICATION NUMBER: US/09/758,449  
;; CURRENT FILING DATE: 2001-01-11  
;; PRIOR APPLICATION NUMBER: 60/179,065  
;; PRIOR FILING DATE: 2000-01-31  
;; PRIOR APPLICATION NUMBER: 60/180,628  
;; PRIOR FILING DATE: 2000-02-04  
;; NUMBER OF SEQ ID NOS: 1478  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 1158  
;; LENGTH: 341  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: SITE  
;; LOCATION: (215)  
;; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-758-449-1158

Query Match 92.0%; Score 674; DB 21; Length 341;  
Best Local Similarity 99.2%; Pred. No. 3.5e-57;  
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60  
Db 219 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 278

QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 120  
Db 279 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 338

QY 121 Q 121  
Db 339 Q 339

RESULT 15  
US-09-760-443-1328  
Sequence 1328, Application US/09760443  
GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PZ12  
CURRENT APPLICATION NUMBER: US/09/760,443  
CURRENT FILING DATE: 2001-01-16  
Prior application data removed - refer to PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2164  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1328  
LENGTH: 341  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: SITE  
LOCATION: (215)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-760-443-1328

Query Match 92.0%; Score 674; DB 21; Length 341;  
Best Local Similarity 99.2%; Pred. No. 3.5e-57;  
Matches 120; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 60  
Db 219 CEPLAPELLGMDCTHPGNSFSSQCAFSCSEGTNLGIEETTCGPGNWSPEPTCOV 278

QY 61 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 120  
Db 279 IQCEPLASAPDLGIMNCSHPLASFSTSACTFTICSEGTLLGKKTICSSGIMNSPSPIC 338

OY 121 Q 121  
DB 339 Q 339

```
RESULT 16
US-09-119-209-4
; Sequence 4, Application US/09119209
; GENERAL INFORMATION:
; APPLICANT: LASKY, LAURENCE A.
; APPLICANT: STACHELL, SCOTT E.
; APPLICANT: ROSEN, STEVEN D.
; APPLICANT: SINGER, MARK S.
; APPLICANT: YEDNOCK, TED A.
; TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 1 DNA Way
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WinPatIn (Genentech)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/119,209
; FILING DATE: 20-Jul-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/513278
; FILING DATE: 10-AUG-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/059027
; FILING DATE: 6-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/786149
; FILING DATE: 31-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/315015
; FILING DATE: 23-FEB-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Love, Richard B.
; REGISTRATION NUMBER: 34,659
; REFERENCE/DOCKET NUMBER: P0565D1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650/225-5530
; TELEFAX: 650/952-9881
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 372 amino acids
; TYPE: Amino Acid
; TOPOLOGY: Linear
; US-09-119-209-4

Query Match          70.7%; Score 518; DB 15; Length 372;
Best Local Similarity 68.2%; Pred. No. 6, 6e-42;
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;
```

```
OY 1 CEPLAEPGLTMDCTHPFGNFSFSCAFSCSEGTNLGTIEETTCGPFGNMSSPEPTCOY 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 197 CEPLAEPGLTMDCTHPFGNFSFSCAFSCSEGTNLGTIEETTCGPFGNMSSPEPTCOY 256
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 ICEPLASAPDLGTMCNCSHPLASFTSACTFTCSGTELTIGKKTKTCSSGTSNPSPTC 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 257 VOCEPLAEPGLTMDCTHPFGNFSFSCAFSCSEGTNLGTIEETTCGPFGNMSSPEPTC 316
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

DB 317 QETNRSFSKIKE 328

```
RESULT 17
PCT-US94-09395-4
; Sequence 4, Application PC/TUS9409395
; GENERAL INFORMATION:
; APPLICANT: Board of Regents of the University of Oklahoma
; TITLE OF INVENTION: Expression Control Sequences of the P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: GA
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/09395
; FILING DATE: 19-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404)-815-6508
; TELEFAX: (404)-815-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 830 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US94-09395-4
```

Query Match 43.0%; Score 315; DB 1; Length 830;  
Best Local Similarity 43.9%; Pred. No. 1, 1e-21;  
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```
OY 1 CEPLAEPGLTMDCTHPFGNFSFSCAFSCSEGTNLGTIEETTCGPFGNMSSPEPTCOY 60
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 324 CQLEAPSEGTMDCTHPFGNFSFSCAFSCSEGTNLGTIEETTCGPFGNMSSPEPTCOY 383
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
OY 61 ICEPLASAPDLGTMCNCSHPLASFTSACTFTCSGTELTIGKKTKTCSSGTSNPSPTC 120
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
DB 384 ICEPLESPVHGSMDCSPSLRAFQYDTNCSFCAEGFMGLGADIVCDNLGQATAPAYC 443
|:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

OY 121 QKL 123  
DB 444 QAL 446

```
RESULT 18
US-08-449-687B-4
; Sequence 4, Application US/08449687B
; GENERAL INFORMATION:
; APPLICANT: McEver, Rodger P.
; APPLICANT: Pan, Junliang
; TITLE OF INVENTION: Expression Control Sequences of the
; TITLE OF INVENTION: P-Selectin Gene
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
```

```

STATE: GA
COUNTRY: USA
ZIP: 30309-3450
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/449,687B
FILING DATE: 24-MAY-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/110,158
FILING DATE: 20-AUG-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/320,408
FILING DATE: 08-MAR-1989
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
TELECOMMUNICATION INFORMATION:
TELEPHONE: (404)873-8794
TELEFAX: (404)873-8795
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 830 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-449-687B-4

```

```

Query Match 43.0%; Score 315; DB 8; Length 830;
Best Local Similarity 43.9%; Pred. No. 1.1e-21;
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```

```

QY 1 CEPELAPELGTMDCVTHPRGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 CQHLAPSEGTMDCVHPLTAFAVGSCKFCQPGYRVRGLMLRCIDSGHWSAPLPTCEA 383

QY 61 IQCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTLLGKKTKICSSGIMNSPSPIC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 ISCEPLSPVHGSMDCSPSLRAFOYDTMCSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 443

QY 121 OKL 123
Db 444 QAL 446

```

```

RESULT 19
US-10-020-141-10
; Sequence 10, Application US/10020141.
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Jeanette
; APPLICANT: Ableson, Allen
; TITLE OF INVENTION: DIAGNOSIS AND TREATMENT OF VASCULAR DISEASE
; FILE REFERENCE: MMI-002
; CURRENT APPLICATION NUMBER: US/10/020,141
; CURRENT FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/313,097
; PRIOR FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: US 60/327,485
; PRIOR FILING DATE: 2001-10-05
; NUMBER OF SEQ ID NOS: 21
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 830
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-020-141-10

```

```

Query Match 43.0%; Score 315; DB 24; Length 830;
Best Local Similarity 43.9%; Pred. No. 1.1e-21;
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

```

```

QY 1 CEPELAPELGTMDCVTHPRGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 324 CQHLAPSEGTMDCVHPLTAFAVGSCKFCQPGYRVRGLMLRCIDSGHWSAPLPTCEA 383

QY 61 IQCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTLLGKKTKICSSGIMNSPSPIC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 384 ISCEPLSPVHGSMDCSPSLRAFOYDTMCSFRCAEGFMLRGADIVRCDNLGQWTAPAPVC 443

QY 121 OKL 123
Db 444 QAL 446

```

```

RESULT 20
US-60-207-315-467
; Sequence 467, Application US/60207315
; GENERAL INFORMATION:
; APPLICANT: Beasley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PROTEASE PROTEINS, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000601
; CURRENT APPLICATION NUMBER: US/60/207,315
; CURRENT FILING DATE: 2000-05-30
; NUMBER OF SEQ ID NOS: 528
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 467
; LENGTH: 700
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: (1)...(700)
; OTHER INFORMATION: Xaa = Any Amino Acid
US-60-207-315-467

```

```

Query Match 41.7%; Score 306; DB 26; Length 700;
Best Local Similarity 41.3%; Pred. No. 6.7e-21;
Matches 50; Conservative 23; Mismatches 48; Indels 0; Gaps 0;

```

```

QY 1 CEPELAPELGTMDCVTHPRGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 192 CGELELPQHVLMNCSHPLGNFSFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 251

QY 61 IQCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTLLGKKTKICSSGIMNSPSPIC 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 252 AQCPPLKIPERGNMTCILSKAKAFQHOSSCSFSCHEGALVGVYQCTAGAWTAPAPVC 311

QY 121 Q 121
Db 312 K 312

```

```

RESULT 21
PCT-US99-28965-19
; Sequence 19, Application PC/TUS9928965
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; TITLE OF INVENTION: alpha-INDUCED EXPRESSION OF CELL ADHESION MOLECULES
; FILE REFERENCE: ISPH-0424
; CURRENT APPLICATION NUMBER: PCT/US99/28965
; CURRENT FILING DATE: 1999-12-08
; EARLIER APPLICATION NUMBER: US 09/209,668
; EARLIER FILING DATE: 1998-12-10

```



QY 121 Q 121  
Db 300 K 300

RESULT 24  
US-09-266-091-2  
; Sequence 2, Application US/09266091  
; GENERAL INFORMATION:  
; APPLICANT: Klimuk, Sandra K  
; APPLICANT: Klimuk, Sean C  
; APPLICANT: Scherret, Peter  
; APPLICANT: Hope, Michael J.  
; APPLICANT: Bennett, C. Frank  
; TITLE OF INVENTION: ENHANCED EFFICACY OF LIPOSOMAL ANTISENSE THERAPY  
; FILE REFERENCE: ISPH-0342  
; CURRENT APPLICATION NUMBER: US/09/266,091  
; CURRENT FILING DATE: 1999-03-10  
; PRIOR APPLICATION NUMBER: 08/657,753  
; PRIOR FILING DATE: 1996-05-30  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-266-091-2

Query Match 40.4%; Score 296; DB 16; Length 610;  
Best Local Similarity 41.3%; Pred. No. 5.5e-20;  
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1  
Db 180  
QY 61  
Db 240  
QY 121 Q 121  
Db 300 K 300

RESULT 25  
US-09-266-091A-2  
; Sequence 2, Application US/09266091A  
; GENERAL INFORMATION:  
; APPLICANT: Klimuk, Sandra K.  
; APPLICANT: Semple, Sean C.  
; APPLICANT: Scherret, Peter  
; APPLICANT: Hope, Michael J.  
; TITLE OF INVENTION: Enhanced Efficacy of Liposomal Antisense  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Townsend and Townsend and Crew LLP  
; STREET: Two Embarcadero Center, Eighth Floor  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94111-3834  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/266,091A  
; FILING DATE: 10-Mar-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/657,753  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Kezer, William B.  
REGISTRATION NUMBER: 37,369  
REFERENCE/DOCKET NUMBER: 16303-003600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 610 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-266-091A-2

Query Match 40.4%; Score 296; DB 16; Length 610;  
Best Local Similarity 41.3%; Pred. No. 5.5e-20;  
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1  
Db 180  
QY 61  
Db 240  
QY 121 Q 121  
Db 300 K 300

RESULT 26  
US-09-784-356-122  
; Sequence 122, Application US/09784356  
; GENERAL INFORMATION:  
; APPLICANT: Murray, Richard  
; APPLICANT: Glynn, Richard  
; APPLICANT: Watson, Susan R.  
; APPLICANT: EOS Biotechnology, Inc.  
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,  
; TITLE OF INVENTION: Compositions and Methods of Screening for Angiogenesis  
; FILE REFERENCE: 018501-00071005  
; CURRENT APPLICATION NUMBER: US/09/784,356  
; CURRENT FILING DATE: 2001-02-14  
; PRIOR APPLICATION NUMBER: US 60/148,425  
; PRIOR FILING DATE: 1999-08-11  
; PRIOR APPLICATION NUMBER: US 09/637,977  
; PRIOR FILING DATE: 2000-08-11  
; NUMBER OF SEQ ID NOS: 135  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 122  
; LENGTH: 610  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-784-356-122

Query Match 40.4%; Score 296; DB 21; Length 610;  
Best Local Similarity 41.3%; Pred. No. 5.5e-20;  
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1  
Db 180  
QY 61  
Db 240  
QY 121 Q 121  
Db 300 K 300



```
Db 240 VECDAVTNPANGFVECFONPGSFPMNTTCTFDCEEGFELMGASLOCTSSGNNDNKPTC 299
QY 121 Q 121
Db 300 K 300

RESULT 27
US-09-802-640-36
; Sequence 36, Application US/09802640
; GENERAL INFORMATION:
; APPLICANT: Braun, Andreas
; APPLICANT: Bonsal Aruna
; APPLICANT: Kleya Patrick
; TITLE OF INVENTION: GENES AND POLYMORPHISMS ASSOCIATED WITH
; FILE REFERENCE: 24736-2048
; CURRENT APPLICATION NUMBER: US/09/802,640
; CURRENT FILING DATE: 2001-03-09
; NUMBER OF SEQ ID NOS: 122
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 36
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-802-640-36

Query Match 40.4%; Score 296; DB 22; Length 610;
Best Local Similarity 41.3%; Pred. No. 5.5e-20;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1CEPLAPELLGTMDCTHPFNFSFSSQCAFSCSEGTNLGTGIEERTCGPFGWSSPEPTCOY 60
| 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 180 CTALSPERHSLVCSHPIDNFSTNSSCSISCDRGYLPSSMETWQCMSSGWSAPIPACNV 239
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKTCSSGATWSNPSPIC 120
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 240 VECDAVTNPANGFVECFONPGSFPMNTTCTFDCEEGFELMGASLOCTSSGNNDNKPTC 299
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 121 Q 121
Db 300 K 300

RESULT 28
US-09-857-670-19
; Sequence 19, Application US/09857670
; GENERAL INFORMATION:
; APPLICANT: Monia, Brett P.
; APPLICANT: Xu, Xiaoxing S.
; APPLICANT: Isis Pharmaceuticals, Inc.
; TITLE OF INVENTION: METHODS OF MODULATING TUMOR NECROSIS FACTOR
; FILE REFERENCE: ISPH-0424
; CURRENT APPLICATION NUMBER: US/09/857,670
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: US 09/209,668
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 19
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-857-670-19

Query Match 40.4%; Score 296; DB 22; Length 610;
Best Local Similarity 41.3%; Pred. No. 5.5e-20;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1CEPLAPELLGTMDCTHPFNFSFSSQCAFSCSEGTNLGTGIEERTCGPFGWSSPEPTCOY 60
| 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 180 CTALSPERHSLVCSHPIDNFSTNSSCSISCDRGYLPSSMETWQCMSSGWSAPIPACNV 239
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKTCSSGATWSNPSPIC 120
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 240 VECDAVTNPANGFVECFONPGSFPMNTTCTFDCEEGFELMGASLOCTSSGNNDNKPTC 299
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 121 Q 121
Db 300 K 300

RESULT 29
US-10-021-660-122
; Sequence 122, Application US/10021660
; GENERAL INFORMATION:
; APPLICANT: Murray, Richard
; APPLICANT: Glynn, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: EOS Biotechnology, Inc.
; TITLE OF INVENTION: Novel Methods of Diagnosis of Angiogenesis,
; FILE REFERENCE: 018501-000710US
; CURRENT APPLICATION NUMBER: US/10/021,660
; CURRENT FILING DATE: 2001-12-06
; PRIOR APPLICATION NUMBER: US/09/784,356
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: US 09/637,977
; PRIOR FILING DATE: 2000-08-11
; NUMBER OF SEQ ID NOS: 135
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 122
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-021-660-122

Query Match 40.4%; Score 296; DB 24; Length 610;
Best Local Similarity 41.3%; Pred. No. 5.5e-20;
Matches 50; Conservative 21; Mismatches 50; Indels 0; Gaps 0;

QY 1 1CEPLAPELLGTMDCTHPFNFSFSSQCAFSCSEGTNLGTGIEERTCGPFGWSSPEPTCOY 60
| 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 180 CTALSPERHSLVCSHPIDNFSTNSSCSISCDRGYLPSSMETWQCMSSGWSAPIPACNV 239
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 61 IOCEPLSAPDLGIMNCSHPLASFSTACTFICSEGTGELIGKKTKTCSSGATWSNPSPIC 120
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11
Db 240 VECDAVTNPANGFVECFONPGSFPMNTTCTFDCEEGFELMGASLOCTSSGNNDNKPTC 299
::1: 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11 11:11

QY 121 Q 121
Db 300 K 300

RESULT 30
PCT-US01-01239-1145
; Sequence 1145, Application PC/TUS0101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1145
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

Query Match	36.6%	Score 268;	DB 21;	Length 355;
Best Local Similarity	39.7%	Pred. No. 1.6e-17;		
Matches	48;	Conservative	17;	Mismatches 56;
			Indels	0;
			Gaps	0;

Query Match	34.8%;	Score 255;	DB 21;	Length 309;
Best Local Similarity	38.8%;	Pred. No. 2.6e-16;		
Matches 47; Conservative	16;	Mismatches 58;	Indels 0;	Gaps 0;



```

? SOFTWARE:PatentIn Ver. 2.0
? SEQ ID NO 920
? LENGTH: 207
? TYPE: PRT
? ORGANISM: Homo sapiens
? FEATURE:
? NAME/KEY: SITE
? LOCATION: (151)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (185)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (191)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (194)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
? NAME/KEY: SITE
? LOCATION: (198)
? OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-760-498-920

Query Match          32.2%; Score 236; DB 21; Length 207;
Best Local Similarity 37.4%; Pred. No. 1.2e-14;
Matches 46; Conservative 17; Mismatches 54; Indels 6; Gaps 3

QY      1 CEPLAEPELGMCTH-PFGNFSFSSOCAPSCSEGTNLGIEETTCGFPGMSSPEPTCQ 59
Db       23 CDVAHQPKGLVRAMHPISGEFTFKSSCAPSCSEGFLHGSTOLCTSGQGWTVEVPSCQ 82
         :.:|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
QY      60 VIOCEPLSAPDLGIMNC-S-HPLASFSTACTFTICSEGTELIGKKRITCESSGIWNSPSP 118
Db       83 VVKCSSLAIVPGKIMSCSEPV---FGVYCKFACEGWTLNGSAARFCATGHMSGLLP 138
         |:|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|.|
QY      119 ICQ 121
         |:
Db       139 TCE 141

RESULT 39
PCT-US01-01239-1688
; Sequence 1688, Application PC/TUSO101239
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc., et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT213PCT
; CURRENT APPLICATION NUMBER: PCT/US01/01239
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2318
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1688
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (31)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (76)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (127)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (135)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
PCT-US01-01239-1688
```



**This Page blank (uspto)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 7, 2002, 10:14:43 : Search time 28.63 Seconds

(without alignments)  
112,615 Million cell updates/sec

Title: US-09-119-209-2\_COPY\_197\_328

Perfect score: 733  
Sequence: 1 CEPLEAPELGTMDCTHPRGN.....WSNPSPICKKDKSRSMIKE 132Scoring table:  
BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued\_Patents\_AA : \*

- 1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	733	100.0	372	2	US-08-513-278-2
2	733	100.0	372	6	5514582-2
3	713	97.3	385	1	US-08-340-539A-2
4	713	97.3	385	2	US-08-461-592B-2
5	518	70.7	372	2	US-08-513-278-4
6	518	70.7	372	6	5514582-4
7	487	66.4	126	6	5514582-31
8	315	43.0	830	1	US-08-110-158-4
9	314	42.8	830	5	PCT-US91-05059-2
10	305	41.4	574	6	5378464-3
11	303.5	41.4	830	6	5378464-2
12	296	40.4	610	1	US-08-365-470-3
13	296	40.4	610	3	US-09-209-668-19
14	296	40.4	610	4	US-09-009-490A-89
15	296	40.4	610	6	5217870-2
16	245	33.4	484	2	US-08-252-493C-9
17	245	33.4	484	3	US-09-276-197-9
18	165.5	22.6	240	3	US-08-824-692-23
19	163.5	22.3	127	6	5514582-32
20	163.5	22.3	1466	6	5256642-6
21	163.5	22.3	1466	6	5472939-6
22	163.5	22.3	1537	6	5256642-5
23	163.5	22.3	1537	6	5472939-5
24	163.5	22.3	1847	6	5256642-10
25	163.5	22.3	1847	6	5472939-10
26	163.5	22.3	2039	6	5256642-2
27	163.5	22.3	2039	6	5472939-2

28	162	22.1	128	6	5514582-33	Patent No. 5514582
29	153	20.9	62	1	US-08-202-047-20	Sequence 20, Appl
30	153	20.9	62	4	US-08-964-690-20	Sequence 20, Appl
31	152.5	20.8	363	4	US-08-961-234B-2	Sequence 2, Appl1
32	149	20.3	216	6	US-08-824-692-24	Sequence 24, Appl
33	148.5	20.3	126	6	5514582-35	Patent No. 5514582
34	140.5	19.2	265	2	US-08-177-109A-57	Sequence 57, Appl
35	140.5	19.2	265	2	US-08-687-706-57	Sequence 57, Appl
36	140.5	19.2	764	2	US-08-177-109A-2	Sequence 2, Appl1
37	140.5	19.2	764	2	US-08-687-706-2	Sequence 2, Appl1
38	139	19.0	126	6	5514582-43	Patent No. 5514582
39	139	19.0	263	1	US-07-906-983-2	Sequence 2, Appl1
40	139	19.0	323	2	US-08-435-149-2	Sequence 2, Appl1
41	139	19.0	324	1	US-08-310-416A-14	Sequence 14, Appl
42	139	19.0	324	2	US-08-888-171-14	Sequence 14, Appl
43	139	19.0	577	2	US-08-435-149-3	Sequence 3, Appl1
44	139	19.0	611	4	US-09-475-460A-32	Sequence 32, Appl
45	138	18.8	197	2	US-08-356-361-27	Sequence 27, Appl

## ALIGNMENTS

RESULT 1  
Sequence 2, Application US/08513278  
Patent No. 5840844  
GENERAL INFORMATION:  
APPLICANT: LASKY, LAURENCE A.  
APPLICANT: STACHELL, SCOTT E.  
APPLICANT: ROSEN, STEVEN D.  
APPLICANT: SINGER, MARK S.  
APPLICANT: YEDNOCK, TED A.  
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/513,278  
FILING DATE: 10-AUG-1995  
CLASSIFICATION: 5530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/059027  
FILING DATE: 06-MAY-1993  
APPLICATION NUMBER: 07/786149  
FILING DATE: 31-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/315015  
FILING DATE: 23-FEB-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Dreger, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 56501C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-513-278-2

	Query Match	100.0%	Score 733	DB 2	length 372	
	Best Local Similarity	100.0%	Pred. NO. 4.1e-71			
	Matches 132	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1	CEPLEAPELGTMDCHHPGNGFSSSCAFCSCSEGTNLGIEETCGPFGNMSSPEPTCOV	60			
Db	197	CEPLEAPELGTMDCHHPGNGFSSSCAFCSCSEGTNLGIEETCGPFGNMSSPEPTCOV	256			
Qy	61	IQCEPLSAPDGLIMNCSHPLASFSTTSACTFLICSECTELIGKKKTIICSSGIWNSDPIC	120			
Db	257	IQCEPLSAPDGLIMNCSHPLASFSTTSACTFLICSECTELIGKKKTIICSSGIWNSDPIC	316			
Qy	121	OKLDRFSFMIRE 132				
Db	317	OKLDRFSFMIRE 328				

RESULT 2  
5514582-2  
Patent NO. 5514582  
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 43  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,670  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 986,931  
FILING DATE: 08-DEC-1992  
APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989  
SEQ ID NO: 2:  
LENGTH: 372  
5514582-2

Query Match	100.0%	Score 733:	DB 6:	Length 372:
Best Local Similarity	100.0%	Pred. No. 4,1e-71:		
Matches 132:	Conservative 0:	Mismatches 0:	Indels 0:	Gaps 0:
QY	1	CEPLEAPELGTMDCTHPPGNFSFSQCAFSQSEGTNLGIEETTCGPPGNMSSPEPTQY	60	
Db	197	CEPLEAPELGTMDCTHPPGNFSFSQCAFSQSEGTNLGIEETTCGPPGNMSSPEPTQY	256	
QY	61	IOCEPLAPDGLGIMNCSPHPLASFSTSACTPCSEGTELIGKKTICSSGIWNSPSPIC	120	
Db	257	IOCEPLAPDGLGIMNCSPHPLASFSTSACTPCSEGTELIGKKTICSSGIWNSPSPIC	316	
QY	121	OKLDKSFMSIKE	132	
Db	317	OKLDKSFMSIKE	328	

RESULT 3  
US-08-340-539A-2  
Sequence 2, Application US/08340539A  
Patent No. 5808025  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
TITLE OF INVENTION: CHIMERIC SELECTINS AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 28  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas

```

: CITY: New York
: STATE: New York
: COUNTRY: USA
: ZIP: 10020
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentlin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/340,539A
: FILING DATE: 16-NOV-1994
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 08/008,459
: FILING DATE: 25-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Gunnison, Jane
: REGISTRATION NUMBER: 38,479
: REFERENCE/DOCKET NUMBER: CG-104 CON
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 212-596-9000
: TELEFAX: 212-596-9090
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 385 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-340-539A-2

```

Query Match Similarity	97.3%;	Score 713;	DB 1;	Length 385;
Best Local Similarity	97.0%;	Pred. No. 6e-69;		
Matches 128;	Conservative 2;	Mismatches 2;	Indels 0;	Gaps 0;
OY	1	CEPLEAPELGTMDCTHHPGNEFSFSQCAFCSCSEGTNLGIEETTCGPPGNMWSPEPTCOV	60	
		:		
Db	210	CEPLEAPELGTMDCTHHPGNEFNFNSQCAFSCSEGTNLGIEETTCGPPGNMWSPEPTCOV	269	
OY	61	IQCEPLASAPDLGIMNCSHPLASFSFTSACTFTCSGTETLIGKKKTTICSSGIWNSPSPIC	120	
Db	270	IQCEPLASAPDLGIMNCSHPLASFSFTSACTFTCSGTETLIGKKKTTICSSGIWNSPSPIC	329	
OY	121	OKLDSFSMIKE	132	
Db	330	OKLDSFSMIKE	341	

RESULT 4  
US-08-461-592B-2  
Sequence 2, Application US/08461592B  
Patent No. 5834425  
GENERAL INFORMATION:  
APPLICANT: Tedder, Thomas F.  
APPLICANT: Kansas, Geoffrey S.  
TITLE OF INVENTION: CHEMERIC SELECTING AS SIMULTANEOUS  
TITLE OF INVENTION: BLOCKING AGENTS FOR COMPONENT SELECTIN FUNCTION  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Weingarten, Schurgin, Gagnebin & Hayes  
STREET: Ten Post Office Square  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/461,592B



FILING DATE: 31-OCT-1991  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/340,539  
FILING DATE: 16-NOV-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/008,459  
FILING DATE: 25-JAN-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: James F. Haley, Jr.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: CG-104  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 596-9000  
TELEFAX: (212) 596-9090  
TELEX: 14-8367  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 385 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-461-5928-2

Query Match 97.3%; Score 713; DB 2; Length 385;  
Best Local Similarity 97.0%; Pred. No. 6e-69;  
Matches 128; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60  
DB 210 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 269  
QY 61 IOCEPLAPDLGINCSHPLASFSTACTFTICSEGTTELIGKKTICSSGIMSNPSPIC 120  
DB 270 IOCEPLAPDLGINCSHPLASFSTACTFTICSEGTTELIGKKTICSSGIMSNPSPIC 329  
QY 121 OKLDSFSMIKE 132  
DB 330 OKLDSFSMIKE 341

RESULT 5  
US-08-513-278-4  
Sequence 4, Application US/08513278  
Patent No. 5840844  
GENERAL INFORMATION:  
APPLICANT: LASKY, LAURENCE A.  
APPLICANT: STACHELL, SCOTT E.  
APPLICANT: ROSEN, STEVEN D.  
APPLICANT: SINGER, MARK S.  
APPLICANT: YEDNICK, TED A.  
TITLE OF INVENTION: LYMPHOCYTE HOMING RECEPTORS  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/513,278  
FILING DATE: 10-AUG-1995  
CLASSIFICATION: 5530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/059027  
FILING DATE: 06-MAY-1993

APPLICATION NUMBER: 07/786149  
FILING DATE: 31-OCT-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/315015  
FILING DATE: 23-FEB-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Dregier, Ginger R.  
REGISTRATION NUMBER: 33,055  
REFERENCE/DOCKET NUMBER: 565D1C1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-3216  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 372 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-513-278-4

Query Match 70.7%; Score 518; DB 2; Length 372;  
Best Local Similarity 68.2%; Pred. No. 5.4e-48;  
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60  
DB 197 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 256  
QY 61 IOCEPLAPDLGINCSHPLASFSTACTFTICSEGTTELIGKKTICSSGIMSNPSPIC 120  
DB 257 VOCEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 316  
QY 121 OKLDSFSMIKE 132  
DB 317 QETNRSFSMIKE 328

RESULT 6  
5514582-4  
Patent No. 5514582  
APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
IMMUNOGLOBULINS  
NUMBER OF SEQUENCES: 43  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,670  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 986,931  
FILING DATE: 08-DEC-1992  
APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989  
SEQ ID NO: 4  
LENGTH: 372  
5514582-4

Query Match 70.7%; Score 518; DB 6; Length 372;  
Best Local Similarity 68.2%; Pred. No. 5.4e-48;  
Matches 90; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

QY 1 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 60  
DB 197 CEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 256  
QY 61 IOCEPLAPDLGINCSHPLASFSTACTFTICSEGTTELIGKKTICSSGIMSNPSPIC 120  
DB 257 VOCEPLAELGTMDCIHPGNFSSQCAFSCSEGTNLGIEETTCGPGNMSSPEPTCOV 316

QY 121 OKLKSFSMIKE 132  
1 : : : : :  
Db 317 QETNRSFSKIKE 328

RESULT 7  
5514582-31  
; APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
; TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
; IMMUNOCLOBULIN

NUMBER OF SEQUENCES: 43  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/185,670  
FILING DATE: 21-JAN-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 986,931  
FILING DATE: 08-DEC-1992  
APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989

; SEQ ID NO:31:  
LENGTH: 126  
5514582-31

Query Match 66.4%; Score 487; DB 6; Length 126;  
Best Local Similarity 66.7%; Pred. No. 3e-45;  
Matches 84; Conservative 13; Mismatches 29; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPFGNFSFSSQCAFSCSEGTNLGTIEETGCGPGNMSPEPTCOV 60  
1 : : : : :  
Db 1 CEPLAPELGTMDCTHPFGNFSFSSQCAFSCSEGTNLGTIEETGCGPGNMSPEPTCOV 60

QY 61 IOCEPLAPDGIIMNCSPHPLASFSTACTFICSGTELLGKKTICSSGIMWNSPIC 120  
1 : : : : :  
Db 61 VOCEPLAPELGTMDCTHPFGNFSFSSQCAFSCSEGTNLGTIEETGCGPGNMSPEPTCOV 120

QY 121 OKLKS 126  
1 : : : : :  
Db 121 QETNRS 126

RESULT 8  
US-08-110-158-4  
; Sequence 4, Application US/08110158

; Patent No. 5605821  
; GENERAL INFORMATION:  
; APPLICANT: McEever, Rodger P.  
; APPLICANT: Pan, Junliang  
; TITLE OF INVENTION: Expression Control Sequences of the  
; TITLE OF INVENTION: P-Selectin Gene  
; NUMBER OF SEQUENCES: 17  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Patrea L. Pabst  
; STREET: 1100 Peachtree Street, Suite 2800  
; CITY: Atlanta  
; STATE: GA  
; COUNTRY: USA  
; ZIP: 30309-4530  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/110,158  
; FILING DATE: 19930820  
; CLASSIFICATION: 424

; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/320,408  
; FILING DATE: 08-MAR-1989  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (404)-815-6508  
; TELEFAX: (404)-815-6555  
; INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:  
LENGTH: 830 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-08-110-158-4

Query Match 43.0%; Score 315; DB 1; Length 830;  
Best Local Similarity 43.9%; Pred. No. 9.9e-26;  
Matches 54; Conservative 21; Mismatches 48; Indels 0; Gaps 0;

QY 1 CEPLAPELGTMDCTHPFGNFSFSSQCAFSCSEGTNLGTIEETGCGPGNMSPEPTCOV 60  
1 : : : : :  
Db 324 COHLEAPSEGTMDCTHPFGNFSFSSQCAFSCSEGTNLGTIEETGCGPGNMSPEPTCOV 383

QY 61 IOCEPLAPDGIIMNCSPHPLASFSTACTFICSGTELLGKKTICSSGIMWNSPIC 120  
1 : : : : :  
Db 384 ISCEPLSFVHSGMDSPSLRAFYDTNCSFRCAGFMLRGADIVRCNDLQOWTAPAPVC 443

QY 121 OKL 123  
1 : : : : :  
Db 444 QAL 446

RESULT 9  
PCT-US91-05059-2  
; Sequence 2, Application PC/US9105059

; GENERAL INFORMATION:  
; APPLICANT: Regents of the Board of the University of  
; APPLICANT: Oklahoma  
; TITLE OF INVENTION: Functionally Active Selectin-Derived  
; TITLE OF INVENTION: Peptides  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Kilpatrick & Cody  
; STREET: 100 Peachtree Street, Suite 3100  
; CITY: Atlanta  
; STATE: Georgia  
; COUNTRY: US  
; ZIP: 30303  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US91/05059  
; FILING DATE: 19910717  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/320408  
; FILING DATE: 08-MAR-1989  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/554199  
; FILING DATE: 17-JUL-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pabst, Patrea L.  
; REGISTRATION NUMBER: 31,284  
; REFERENCE/DOCKET NUMBER: OMRP110CIP  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 404-572-6508

TELEFAX: 404-572-6555	NAME/KEY: Disulfide-bond
INFORMATION FOR SEQ ID NO: 2:	LOCATION: 461..474
SEQUENCE CHARACTERISTICS:	FEATURE:
LENGTH: 830 amino acids	NAME/KEY: Disulfide-bond
TYPE: AMINO ACID	LOCATION: 478..505
STRANDEDNESS: single	FEATURE:
TOPOLOGY: linear	NAME/KEY: Disulfide-bond
MOLECULE TYPE: peptide	LOCATION: 510..554
HYPOTHETICAL: YES	FEATURE:
ANTI-SENSE: YES	NAME/KEY: Disulfide-bond
FRAGMENT TYPE: N-terminal	LOCATION: 523..536
ORIGINAL SOURCE:	FEATURE:
ORGANISM: Homo sapien	NAME/KEY: Disulfide-bond
TISSUE TYPE: Blood	LOCATION: 540..567
CELL TYPE: Endothelial	FEATURE:
FEATURE:	NAME/KEY: Disulfide-bond
NAME/KEY: Disulfide-bond	LOCATION: 572..616
LOCATION: 4..25	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 60..158	LOCATION: 585..616
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 131..150	LOCATION: 602..629
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 163..174	LOCATION: 642..686
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 168..183	LOCATION: 655..668
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 185..194	LOCATION: 672..699
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 200..244	LOCATION: 704..748
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 213..226	LOCATION: 717..730
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Disulfide-bond
LOCATION: 230..257	LOCATION: 734..761
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Binding-site
LOCATION: 262..306	LOCATION: 54
FEATURE:	OTHER INFORMATION: /note- "Potential asparagine-linked
NAME/KEY: Disulfide-bond	OTHER INFORMATION: glycosylation site"
LOCATION: 275..288	FEATURE:
FEATURE:	NAME/KEY: Binding-site
NAME/KEY: Disulfide-bond	LOCATION: 98
LOCATION: 292..319	OTHER INFORMATION: /note- "Potential asparagine-linked
FEATURE:	OTHER INFORMATION: glycosylation site"
NAME/KEY: Disulfide-bond	FEATURE:
LOCATION: 324..368	NAME/KEY: Binding-site
FEATURE:	LOCATION: 180
NAME/KEY: Disulfide-bond	OTHER INFORMATION: /note- "Potential asparagine-linked
LOCATION: 337..350	OTHER INFORMATION: glycosylation site"
FEATURE:	FEATURE:
NAME/KEY: Disulfide-bond	NAME/KEY: Binding-site
LOCATION: 354..381	LOCATION: 212
FEATURE:	OTHER INFORMATION: /note- "Potential asparagine-linked
NAME/KEY: Disulfide-bond	OTHER INFORMATION: glycosylation site"
LOCATION: 386..430	FEATURE:
FEATURE:	NAME/KEY: Binding-site
NAME/KEY: Disulfide-bond	LOCATION: 219
LOCATION: 399..412	OTHER INFORMATION: /note- "Potential asparagine-linked
FEATURE:	OTHER INFORMATION: glycosylation site"
NAME/KEY: Disulfide-bond	FEATURE:
LOCATION: 416..443	NAME/KEY: Binding-site
FEATURE:	LOCATION: 411
NAME/KEY: Disulfide-bond	OTHER INFORMATION: /note- "Potential asparagine-linked
LOCATION: 448..492	OTHER INFORMATION: glycosylation site"
FEATURE:	FEATURE:
	NAME/KEY: Binding-site

LENGTH: 010 amino  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-365-470-3



119 ICQ 121

Db 360 SCO 362

RESULT 18  
US-08-824-692-23  
; Sequence 23, Application US/08824692  
; Patent No. 6017703

GENERAL INFORMATION:

APPLICANT: Kinders, Robert J.

APPLICANT: Enfield, David L.

APPLICANT: Hass, G. Michael

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR SCREENING

TITLE OF INVENTION: FOR OR MODULATING A TUMOR ASSOCIATED ANTIGEN

NUMBER OF SEQUENCES: 38

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/824,692

FILING DATE: 08-APR-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Sharkey, Richard G.

REGISTRATION NUMBER: 32,629

REFERENCE/DOCKET NUMBER: 130001,404

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 23:

SEQUENCE CHARACTERISTICS:

LENGTH: 240 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

US-08-824-692-23

Query Match 22.6%; Score 165.5; DB 3; Length 240;  
Best Local Similarity 30.3%; Pred. No. 2,5e-10;  
Matches 33; Conservative 24; Mismatches 47; Indels 5; Gaps 4;

QY 19 GN-FSFSQCAFSCSEGTNLTG-IEETTGGPFGNMSSPEPTCOVIOCEPLAPDLG--IM 74

Db 6 GNFFGVKAVYTCNEGYYLLGEINRECDTDG-WTNDIPICEVYVCLPVTAPENKIVS 64

QY 75 NCSHPLASFTSACFTICSEGTTELIGKKTKICSSGINSNSPICQKL 123

Db 65 SAMPDREYHFGQAVRFVCSGKIEGDDEMHCSDDGFWSKKEPKVEI 113

RESULT 19  
5514582-32

Patent No. 5514582

APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.

TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID

IMMUNOGLOBULINS

NUMBER OF SEQUENCES: 43

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/185,670

FILING DATE: 21-JAN-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 986,931

FILING DATE: 08-DEC-1992

APPLICATION NUMBER: 808,122  
FILING DATE: 16-DEC-1991  
APPLICATION NUMBER: 440,625  
FILING DATE: 22-NOV-1989  
APPLICATION NUMBER: 315,015  
FILING DATE: 23-FEB-1989  
SEQ ID NO: 32  
LENGTH: 127  
5514582-32

Query Match 22.3%; Score 163.5; DB 6; Length 127;  
Best Local Similarity 28.8%; Pred. No. 1.9e-10;  
Matches 36; Conservative 25; Mismatches 47; Indels 17; Gaps 6;

QY 14 CTH-----PFGNFS-----FSSQCAFSCSEGTNLTG-IEETTGGPFGNMSSPEPTCOV 60

Db 2 CGHPGDTPTGTFPLTNGVNFYGVKAVYTCNEGYYLLGEINRECDTDG-WTNDIPICEV 60

QY 61 IOCEPLAPDLG--IMNCSHPLASFTSACFTICSEGTTELIGKKTKICSSGINSNSPICQKL 118

Db 61 VKCLPYTAPENKIVSNAMEPDREYHFGQAVRFVCSGKIEGDDEMHCSDDGFWSKKEPKVEI 119

QY 119 ICQKL 123

Db 120 KCVEI 124

RESULT 20  
5256642-6

Patent No. 5256642

APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,

MINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN

H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.

TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT

RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF

USE THEREOF

NUMBER OF SEQUENCES: 30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/588,128

FILING DATE: 24-SEP-1990

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 412,745

FILING DATE: 26-SEP-1989

APPLICATION NUMBER: 332,865

FILING DATE: 03-APR-1989

APPLICATION NUMBER: 176,532

FILING DATE: 01-APR-1988

SEQ ID NO: 6;

LENGTH: 1466  
5256642-6

Query Match 22.3%; Score 163.5; DB 6; Length 1466;  
Best Local Similarity 31.9%; Pred. No. 4e-09;

Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

QY 1 CEPLAPELLGTMDCTHPGNSFSQCAFSCSEGTNLTGIEETTGGPFGN---WSSPEPT 57

Db 955 CGPPEPFNGVY---HINTDQFGSTVNYSCNEGFRILGSPSTYCLVSGNNVYWDKADJ 1011

QY 58 COVIOCEP---LSAPDLGIMNCSHPLASFTSACFTICSEGT-----ELIGKKTKICE 108

Db 1012 CEIISCEPPTTISNGDF---YSNNFTSFHNGTVVYQCHTGPDDGQLFELVGERSIYCT 1067

QY 109 SS---GINSNSPICQKL 125

Db 1068 SKDDQYGVWSPPPRCISTNK 1088

RESULT 21  
5472939-6

Patent No. 5472939

```
APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARKIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138, 825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588, 128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:6:
; LENGTH: 1466
5472939-6
```

```
Query Match 22.3%; Score 163.5; DB 6; Length 1466;
Best Local Similarity 31.9%; Pred. No. 4e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

QY 1 CEPLAPELLGTMCTHFGNFSSQCAFSCSEGTNLGIEETTCGPFGN--WSSPEPT 57
DB 955 CGPPEPFNGMV---HINTDQFGSTVYVSCNCGFRIGSGSTICLVSGNNVTWDKKAPI 1011
QY 58 CVOIOCEP---LSAPDGINMCSHPLASFSTACTFICSEGT-----ELIGKKTICE 108
DB 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGQLFELVGERSIYCT 1067
QY 109 SS-----GIMSNPSPIQCKLKD 125
DB 1068 SKDDGVGVWSSPPRCISTNK 1088
```

```
RESULT 22
5256642-5
; Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARKIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
; RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
; USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5:
; LENGTH: 1537
5256642-5
```

```
Query Match 22.3%; Score 163.5; DB 6; Length 1537;
Best Local Similarity 31.9%; Pred. No. 4.2e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

QY 1 CEPLAPELLGTMCTHFGNFSSQCAFSCSEGTNLGIEETTCGPFGN--WSSPEPT 57
DB 955 CGPPEPFNGMV---HINTDQFGSTVYVSCNCGFRIGSGSTICLVSGNNVTWDKKAPI 1011
```

```
QY 58 CVOIOCEP---LSAPDGINMCSHPLASFSTACTFICSEGT-----ELIGKKTICE 108
DB 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGQLFELVGERSIYCT 1067
QY 109 SS-----GIMSNPSPIQCKLKD 125
DB 1068 SKDDGVGVWSSPPRCISTNK 1088
```

```
RESULT 23
5472939-5
; Patent No. 5472939
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARKIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: METHOD OF TREATING COMPLEMENT
; MEDIATED DISORDERS
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/138, 825
; FILING DATE: 19-OCT-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 588, 128
; FILING DATE: 24-SEP-1990
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:5:
; LENGTH: 1537
5472939-5
```

```
Query Match 22.3%; Score 163.5; DB 6; Length 1537;
Best Local Similarity 31.9%; Pred. No. 4.2e-09;
Matches 45; Conservative 16; Mismatches 57; Indels 23; Gaps 6;

QY 1 CEPLAPELLGTMCTHFGNFSSQCAFSCSEGTNLGIEETTCGPFGN--WSSPEPT 57
DB 955 CGPPEPFNGMV---HINTDQFGSTVYVSCNCGFRIGSGSTICLVSGNNVTWDKKAPI 1011
QY 58 CVOIOCEP---LSAPDGINMCSHPLASFSTACTFICSEGT-----ELIGKKTICE 108
DB 1012 CEIISCEPPTISNGDF---YSNNRTSFHNGTVVYTCGHTGPDGQLFELVGERSIYCT 1067
QY 109 SS-----GIMSNPSPIQCKLKD 125
DB 1068 SKDDGVGVWSSPPRCISTNK 1088
```

```
RESULT 24
5256642-10
; Patent No. 5256642
; APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN
; H.; MARKIDES, SAVVAS; MARSH, HENRY C. JR.
; TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT
; RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF
; USE THEREOF
; NUMBER OF SEQUENCES: 30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/588,128
; FILING DATE: 24-SEP-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 412,745
; FILING DATE: 26-SEP-1989
; APPLICATION NUMBER: 332,865
; FILING DATE: 03-APR-1989
; APPLICATION NUMBER: 176,532
; FILING DATE: 01-APR-1988
; SEQ ID NO:10:
```



LENGTH: 1847  
5256642-10

Query Match	22.3%	Score 163.5	DB 6	Length 1847
Best Local	31.9%	Pred. No. 5.3e+09		
Matches 45	Conservative 16	Mismatches 57	Indels 23	Gaps 6

QY 1 CEPLAPELGTMDC THPEGNFSFSQCAFSCSEGTNLGTGIEETTCGPEGN--WSSPEPT 57

Db 1462 CGPPEPENGW---HINTDQFGSTVNYSCNEGFRLLGSPSTTCLVSGNNVTWDKKAPI 1518

02 58 CQVQCEP---LSAPDLGIMNCSHPLASFSTSACTFICSEGT-----ELIGKKTTICE 108

Db 1519 CEISCEPPTISNGD-----YSNNRTSFHNGTVVYQCHTGPDGEOLELVGERSIYCT 1574

QY 109 SS---GIMSNPSPICQKLDK 125

Db 1575 SKDDQVGWSSPPRCISTNK 1595

RESULT 25  
5472939-10

;Patent No. 5472939

;Patent No. 5472939

APPLICANT: FEARON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG  
; WINNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN  
; H.; MAKRIDES, SAVVAS; MARSH, HENRY C. JR.

```

NUMBER OF SEQUENCES: 30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 05/08/138, 8255
FILING DATE: 19-OCT-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 588, 128
FILING DATE: 24-SEP-1990
APPLICATION NUMBER: 412, 745
FILING DATE: 26-SEP-1989
APPLICATION NUMBER: 337, 865
FILING DATE: 03-APR-1989
APPLICATION NUMBER: 176, 532
FILING DATE: 01-APR-1988
SEQ ID NO.:10
LENGTH: 2006
4472939-10

```

Query Match	22.3%	Score 163.5;	DB 6;	Length 1847;
Best Local Similarity	31.9%	Pred. No. 5.3e-09;		
Matches 45;	Conservative 16;	Mismatches 57;	Indels 23;	Gaps 6;

QY 1 CEPLAPELGTMCTHPFGNFSFSSQCAFSCSEGTNLTGIEETTGGPFGN--WSSPEPT 57

Db 1462 CGPPEPENGWV--HINTDTQFGSTVNYSCNEGFRILGSPSTTCLVSGNNVTWDDKAPI 151

QY 58 CVOICEP--LSAPDLIMNCSHPLASFSFTSACTFICSEGT-----ELIGKKTICE 108

Db 1519 CEISCEPPTISNGD-----YSNNRTSFHNGTVVITYOCHTGPDGGEQLFELVGRSITYCT 1577

QY 109 SS---GWSNPSPICQKLDK 125

RESULT 26

5256642-2

5256642-2

PATENT NO. 3230842  
 APPLICANT: FERON, DOUGLAS T.; KLICKSTEIN, LLOYD B.; WONG,  
 WINNIE W.; CARSON, GERALD R.; CONICIO, MICHAEL F.; IP, STEPHEN  
 H.; MAKRIDES, SAVVAS; MARSH, HENRY C., JR.  
 TITLE OF INVENTION: COMPOSITIONS OF SOLUBLE COMPLEMENT  
 RECEPTOR 1 (CRI) AND A THROMBOLYTIC AGENT, AND THE METHODS OF  
 USE THEREOF

```
; NUMBER OF SEQUENCES: 30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/588,128  
; FILING DATE: 24-SEP-1990  
; PRIORITY INFORMATION STATE:
```

Query Match	22.3%	Score	163.5	DB	6	Length	2039
Best Local Similarity	31.9%	Pred. No.	6e-09				
Matches	45	Conservative	16	Mismatches	57	Indels	23
						Gaps	6

QY 1 CEPLAPELGTMCTHPEGNFSFSQCAFSCEGNTLTGIEETTGGFGN--WSSPEPT 57

Db 1457 CGPPEPENGW---HINTDTQEGSTVNYSCNEGFRLLIGSPSTTCLVSGNNVTWDDKAPI 1513

```

02 58 CQVIOCEP---LSAPDGINCSHPLASFSTSACTFICSEGT-----ELIGKKTICE 108

```

Db 1514 CEIISCEPPTISNGD-----YSNNRTSFHNGTVVTTYQCHTGPDGEQLFELVGRSIIYCT 1569

QY 109 SS---GWSNPSPICQKLDK 125

Db 1570 SKDDQVGWSSPPRCISTNK 1590

## RESULT 27

5472939-2  
5472939-2  
5472939-2

5472939-2  
5472939-2  
5472939-2

APPLICANT: FEARON, DOUGLAS T.; KLINKSTEIN, LLOYD B.; WONG,  
; WANNIE W.; CARSON, GERALD R.; CONCINO, MICHAEL F.; IP, STEPHEN  
; H. MAKRIDES SAVVAS; MARSH HENRY C. TP

```

? NUMBER OF SEQUENCES: 30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/138,825
? FILING DATE: 19-OCT-1993
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 588,128
? FILING DATE: 24-SEP-1990
? APPLICATION NUMBER: 412,745
? FILING DATE: 26-SEP-1989
? APPLICATION NUMBER: 337,865
? FILING DATE: 03-APR-1989
? APPLICATION NUMBER: 176,532
? FILING DATE: 01-APR-1988
? SEQ ID NO:2:
? LENGTH: 2039
5472939-2

```

Query Match	22.3%	Score 163.5	DB 6	Length 2039
Best Local Similarity	31.9%	Pred. No. 6e-09		
Matches 45; Conservative	16;	Mismatches 57;	Indels 23;	Gaps 6;

QY 1 CEPLAEPELGTMCTHPFGNFSFSSQCAFSCSEGTNLTGIEETTCGPEGN--WSSPEPT 57

Db 1457 CGPPPEFNGMV--HINTDTQFGSTVNSCNEGFRLIGSPSTTCLVSGNNVTWDKKAPI 151

QY 58 CQVQCEP--LSAPDLGIMNCSHPLASFSTSACTFICSEGT-----ELIGKKTICE 108

Db 1514 CEIISCEPPTISNGDF---YSNN

QY 109 SS---GIWSNPSPICQKLDK

```

Db      1570 SKDDQGVGWSSPPRCISTNK 1590

RESULT      28
5514582-33
: Patent No. 5514582
: APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.
: TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID
: IMMUNOGLOBULINS
: NUMBER OF SEQUENCES: 43
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/185,670
: FILING DATE: 21-JAN-1994
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 986,931
: FILING DATE: 08-DEC-1992
: APPLICATION NUMBER: 808,122
: FILING DATE: 16-DEC-1991
: APPLICATION NUMBER: 440,625
: FILING DATE: 22-NOV-1989
: APPLICATION NUMBER: 315,015
: FILING DATE: 23-FEB-1989
: SEQ ID NO.: 33:
: LENGTH: 128
5514582-33

```

```

; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30, 223
; REFERENCE/DOCKET NUMBER: 14137-77
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-202-047-20

Query Match 20.9%; Score 153; DB 1; Length 62;
Best local Similarity 42.6%; Pred. No. 1e-09;
Matches 26; Conservative 9; Mismatches 26; Indels 0; Gaps 0;

1 CEELEAPDELGTMDCTHPFNGESFSQCAFCSCSEGTNLTGIEETTCGPGWMSSEPTCOV 60
| : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 CQDLPPNENARVNCSHPGAFRYSQVCSFTNENEGILLITGASVLTCLATGWMNSVPEPCA 60

```

```

Query Match      22.1%   Score 162:   DB 6: Length 128;
Best Local Similarity 28.8%; Pred. No. 2.7e-10;
Matches 36; Conservative 24; Mismatches 49; Indels 16; Gaps 5;

QY          14 CTH-----PFGN-----FSPSSQCAFSCEGTMLTG-IERTTGPFGNMSSPEPTCOV 60
            | : ||| : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db          2 CGHGDGDPFGSFLRVLAVSGQFEFGAKVVYTCSGGVLLLEIDYRECGAGS-WINDIPLCEV 60
            | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY          61 IQCEPLAPDVG--IMNCSHPLASFTTSACTFCISCTELIGKRRITCCESSGIWNSDP 118
            ::||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db          61 VKCLPVTLENGRIIVSGAETDDEYYFGQVVFECNMSGFKIEGHKEIKHCSEGLWSNEKP 120
            | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY          119 ICORL 123
            | : :
Db          121 RCVEI 125

RESULT 29
US-08-2002-047-20
: Sequence 20, Application US/08202047
: Patent No. 5800815
: GENERAL INFORMATION:
: APPLICANT: CHESNOT, Robert W.
: APPLICANT: POLLEY, Margaret J.
: APPLICANT: PAULSON, James C.
: APPLICANT: JONES, S. Tarran
: APPLICANT: SALDANHA, Jose W.
: TITLE OF INVENTION: Antibodies to P-selectin and Their Uses
: NUMBER OF SEQUENCES: 45
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Townsend and Townsend Kourie and Crew
: STREET: One Market Plaza, Stewart Tower, Suite 2000
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94105
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/202,047
: FILING DATE: 25-FEB-1994
: CLASSIFICATION: 424

```

```

Db          61 I 61
RESULT      30
US-08-964-690-20
; Sequence 20, Application US/08964690
; Patent No. 6033667
; GENERAL INFORMATION:
; APPLICANT: CHESNOT, Robert W.
; APPLICANT: POLLEY, Margaret J.
; APPLICANT: PAULSON, James C.
; APPLICANT: JONES, S. Tarran
; APPLICANT: SALDANHA, Jose W.
; APPLICANT: BENDIG, Mary M.
; TITLE OF INVENTION: Antibodies to P-Selectin and Their Uses
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourie and Crew
; STREET: One Market Plaza, Stewart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/964,690
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/202,047
; FILING DATE: 25-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, William M.
; REGISTRATION NUMBER: 30,223
; REFERENCE/DOCKET NUMBER: 14137-77
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-326-2400
; TELEFAX: 415-326-2422
; INFORMATION FOR SEQ ID NO: 20:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 62 amino acids
; TYPE: amino acid

```



```

? APPLICATION NUMBER: 986,9931
? FILING DATE: 08-DEC-1992
? APPLICATION NUMBER: 808,1222
? FILING DATE: 16-DEC-1991
? APPLICATION NUMBER: 440,6255
? FILING DATE: 22-NOV-1986
? APPLICATION NUMBER: 315,0155
? FILING DATE: 23-FEB-1989
? SEQ ID NO.:35:
? LENGTH: 126
5514582-35

```

Query Match	20.3%	Score 148.5	DB 6	Length 126
Best Local Similarity	28.4%	Pred. No. 7.6e-09		
Matches	31	Conservative	21	Mismatches 48
				Indels 9
				Gaps 3

QY 80 L--ASSTFSACTFICSEBTETLCKKRTICSSG---TMSNPSTCQL 123  
 75 LEPVNIQIQAKAVDFVCDDBEFQKGSASATCYVLAGESTLMSNSVPEVCKI 123  
 Db 19 NFSPGQEVFSCPEGYDLDGASMRCTPGQDMSPAAPTCVEXKSDDE----KQGLNGRY 74  
 QY 20 NEFFSSQACFSCSEEGNLIGIETTCGFGFNMSSSEPPCYQICEPLSTADLDINMCSH 79  
 15 NEFFSSQACFSCSEEGNLIGIETTCGFGFNMSSSEPPCYQICEPLSTADLDINMCSH 79

```

; RESULT 34
; US-08-177-109A-57
; Sequence 57, Application US/08177109A
; Patent No. 5869615
; GENERAL INFORMATION:
; APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby
; TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSSE: Patrea L. Pabst
; STREET: 2800 One Atlantic Center
; STREET: 1201 West Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-3450
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/177,109A
; FILING DATE: 03-JAN-1994
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: MU 107
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 873-8794
; TELEFAX: (404) 873-8795
; INFORMATION FOR SEO ID NO.: 57:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 265 amino acids
; type: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-08-177-109A-57

```

Query Match	19.2%;	Score 140.5;	DB 2;	Length 265;	
Best Local Similarity	25.2%;	Pred. No. 1.4e-07;			
Matches	35;	Mismatches 55;	Indels 35;	Gaps 4	
QY	1	CEPLAARLGTMDCTHPFGN-----FSPSSQCAFCSCSECTNLGTGIBETTCGFGN	51		

```

Db      73 CAIHCPR-----PHDENGEYMPRSPDYVNVSDIEIPHCYDGTLLGSANRTQGVNGRW 126
QY      52 SSPEPTCOVIOCEPLASADDLGIMNCSP-----LASTSFTSACTFICESTELIG 102
Db      127 SCQTAIC-----DNGAGYCSPNGIPIGTRKVSQYRLDSVTHMCSRLTLRGS 175
QY      103 KTIICSSGIMSNPSPICQ 121
Db      176 QARTQEGGSMSCITPSCQ 194

```

RESULT 35  
 US-08-687-706-57  
 Sequence 57, Application US/08687706  
 Patent No. 5928892  
 GENERAL INFORMATION:  
 APPLICANT: Dennis E. Hourcade and Teresa J. Ogilvie  
 TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASIS  
 NUMBER OF SEQUENCES: 62  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center  
 STREET: 1201 West Peachtree Street  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/687,706  
 FILING DATE: 26-JUL-1996  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/177,109  
 FILING DATE: 03-JAN-1994  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: WU 107 DIV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (404) 873-8794  
 TELEFAX: (404) 873-8795  
 INFORMATION FOR SEQ. ID NO: 57:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 265 amino acids  
 TYPE: amino acid  
 TOPOLOGY: Linear  
 MOLECULE TYPE: protein  
 HYPOTHEetical: NO  
 US-08-687-706-57

	Query March	19.28;	Score 140.5;	DB 2;	Length 265;	
	Best Local Similarity	25.28;	Pred. No. 1.4e-07;			
	Matches 35;	Conservative 14;	Mismatches 55;	Indels 35;	Gaps 4;	
QY	1 CEPEAELELTGMDCTHPFCN-----FSRSQCACFSCSEGTNLNGIEETTCGPFGNW 51					
Dd	73 CRAIHCR-----PHDFENGWEYWRSPRYNYSDIETHCIDGTLINGSANRHCQVGNRR 126					
QY	52 SSPEPTQVIQCEPLSAPDLGIMNSHP-----LASFSFTSACTFLCSBGTBLIGK 102					
Dd	127 SGQTALC-----DNGAGYCNSNPICPIGRTRKVSGYRLIEDSVTHYSNGLFRGS 175					
QY	103 KKTICESSGIWSNPFIQ 121					
Dd	176 QRRTCGGGWSMGTEPSCQ 194					

RESULT 36  
 US-08-177-109A-2  
 Sequence 2, Application US/08177109A  
 Patent No. 5859615  
 GENERAL INFORMATION:  
 APPLICANT: Dennis E. Hourcade and Teresa J. Ogile  
 TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEINASE  
 NUMBER OF SEQUENCES: 62  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Patrea L. Pabst  
 STREET: 2800 One Atlantic Center  
 STREET: 1201 West Peachtree Street  
 CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/177,109A  
 FILING DATE: 03-JAN-1994  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patrea L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: WU 107  
 TELECOMMUNICATION INFORMATION:

```

? INFORMATION FOR SEQ ID NO: 2:
?     SEQUENCE CHARACTERISTICS:
?         LENGTH: 764 amino acids
?         TYPE: amino acid
?         TOPOLOGY: linear
?     MOLECULE TYPE: protein
?     HYPOTHEetical: NO
? OS-08-177-109A-2

```

	Query Match	19.28:	Score 140.5;	DB 2;	Length 764;
	Best Local Similarity	25.2%;	Pred. No. 5.2e-07;		
	Matches 35; Conservative	14;	Mismatches 55;	Indels 35;	Gaps
OY	1 CEPLAPELLGMDCTHPGN-----FSSQCAFCSCSEGTNLGIEETTCGPNGW	51			
Db	98 CRAIHCP-----PHDENGEIYPFRSPRYVNSDEISHCHDGLTLGSNARTQVNGRW	15			
OY	52 SSPEPTCOVIOCEPLSAPDLGINNCSDP-----LASFSPTSACTFICSESTELIGK	107			
Db	152 SGOTAIAC-----DNAGAGCYSNGEIGIRKRVSGYRLDSVTYHCSRELTLRGS	200			
OY	103 KKTICSESGIWSNPISQIQ	121			
Db	201 QRTCCQEGGSMWGTEPSQQ	219			

US-RESULT 37  
US-08-687-706-2  
Sequence 2, Application US/08687706  
Patent No. 5928692  
GENERAL INFORMATION:  
APPLICANT: Dennis E. Hourcade and Teresa J. Oglesby  
TITLE OF INVENTION: MODIFIED COMPLEMENT PROTEASES  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Patricia L. Pabst  
STREET: 2800 One Atlantic Center  
STREET: 1201 West Peachtree Street

CITY: Atlanta  
 STATE: Georgia  
 COUNTRY: USA  
 ZIP: 30309-3450  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 OPERATING SYSTEM: IBM PC compatible  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/687,706  
 FILING DATE: 26-JUL-1996  
 CLASSIFICATION: 514  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/177,109  
 FILING DATE: 03-JAN-1994  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pabst, Patricia L.  
 REGISTRATION NUMBER: 31,284  
 REFERENCE/DOCKET NUMBER: WU 107 DIV  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (404) 873-8794  
 TELEFAX: (404) 873-8795  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 764 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 JS-08-687-706-2

	Query Match	Similarity	19.2%;	Score	140.5;	DB 2;	length	764;	
	Best Local	Similarity	25.2%;	Pred.	No. 5.2e-07;				
	Matches	Conservative	35;	Mismatches	55;	Indels	35;	Gaps	4
QY	1	CEPLEAPELIGTMDCTHPPGN-----FSPSQAFAFCSEGTNLTGIEETTCGPFNGW	51						
Db	98	CRAIHICPR-----PHDENGEYWPDRSPYANNVDSDEISFYCYDGYTLRGSANNTCOVNGMW	151						
QY	52	SSPEPTCOVIQCEPLAPDLGITMNCSHP-----LASFTSACTPIGSEGTLLICK	102						
Db	152	SGQTALIC-----DNGAGYCNSNPIDPIGTKRVGGQYRLDESVTIHCRCGLTLRGS	200						
QY	103	KKTICESSGIWSNPSPIQO 121							
Db	201	QRRTQEGGSMWGTEPSCQ 219							

RESULT 38  
 5514582-43  
 Patent NO. 5514582  
 APPLICANT: CAPON, DANIEL J.; LASKY, LAURENCE A.  
 TITLE OF INVENTION: RECOMBINANT DNA ENCODING HYBRID  
 IMMUNOGLOBULINS  
 NUMBER OF SEQUENCES: 43  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/185,670  
 FILING DATE: 21-JAN-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 986,931  
 FILING DATE: 08-DEC-1992  
 APPLICATION NUMBER: 808,122  
 FILING DATE: 16-DEC-1991  
 APPLICATION NUMBER: 440,625  
 FILING DATE: 22-NOV-1989  
 APPLICATION NUMBER: 315,015  
 FILING DATE: 23-FEB-1989  
 SEQ ID NO: 43:  
 LENGTH: 126  
 5514582-43



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

Om protein - protein search, using sw model

Run on: September 7, 2002, 10:14:07 ; Search time 72.54 Seconds  
(without alignments)  
202.119 Million cell updates/sec

```

Title:      US-09-119-209-2_COPY_197_328
Perfect score: 733
Sequence:   1 CEPLAEPLGTMDCTHPEGN.....WSNPSPICQKIDKFSMIKE 132

```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

```
Post-processing:  Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries
```

Database :

1:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT *
A:cgenseq.032802.*	
2:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT *
3:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT *
4:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT *
5:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT *
6:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT *
7:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT *
8:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT *
9:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT *
10:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT *
11:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT *
12:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT *
13:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT *
14:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT *
15:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT *
16:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT *
17:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT *
18:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT *
19:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT *
20:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT *
21:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT *
22:	/SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	length	DB	ID	Description
1	733	100.0	369	17	AAR98125	Human lymphocyte c
2	733	100.0	370	17	AAR98126	Human lymphocyte c
3	733	100.0	370	17	AAR98127	Human lymphocyte c
4	733	100.0	371	17	AAR98109	Human lymphocyte c
5	733	100.0	371	17	AAR98110	Human lymphocyte c
6	733	100.0	371	17	AAR98111	Human lymphocyte c
7	733	100.0	371	17	AAR98112	Human lymphocyte c
8	733	100.0	371	17	AAR98113	Human lymphocyte c
9	733	100.0	371	17	AAR98114	Human lymphocyte c
10	733	100.0	371	17	AAR98115	Human lymphocyte c
11	733	100.0	371	17	AAR98116	Human lymphocyte c

12	733	100.0	372	12	AAR12469	Human lymphocyte c
13	733	100.0	372	13	AAR24026	Sequence of human
14	733	100.0	372	14	AAR37960	Human lymphocyte H
15	733	100.0	372	14	AAR38908	HuLHR, Homo sapie
16	733	100.0	372	16	AAR76506	Human LHR, Homo s
17	733	100.0	372	16	AAR83050	Human LHR, Homo s
18	733	100.0	372	17	AAR98106	Human lymphocyte c
19	733	100.0	372	19	AAR37781	Homo sapiens lymph
20	733	100.0	372	19	AAW73264	Human lymphocyte f
21	733	100.0	374	17	AAR98131	Human lymphocyte c
22	733	100.0	374	17	AAR98132	Human lymphocyte c
23	732	99.9	371	17	AAR98122	Human lymphocyte c
24	731	99.7	371	17	AAR98124	Human lymphocyte c
25	730	99.6	371	17	AAR98119	Human lymphocyte c
26	729	99.5	371	17	AAR98117	Human lymphocyte c
27	729	99.5	371	17	AAR98121	Human lymphocyte c
28	728	99.3	372	13	AAR22802	Human lymphocyte f
29	727	99.2	371	17	AAR98118	Human lymphocyte c
30	727	99.2	371	17	AAR98120	Human lymphocyte c
31	727	99.2	371	17	AAR98123	Human lymphocyte c
32	727	99.2	372	22	AAB68134	Amino acid sequenc
33	722.5	98.6	372	17	AAR98133	Human lymphocyte c
34	721.5	98.4	374	17	AAR98134	Human lymphocyte c
35	721.5	98.4	374	17	AAR98135	Human lymphocyte c
36	719	98.1	363	13	AAR22551	T lymphocyte-speci
37	719	98.1	363	13	AAR91443	Human leu8 antigen
38	719	98.1	363	21	AAI96184	Human T-cell speci
39	719	98.1	385	13	AAR20815	T lymphocyte-speci
40	719	98.1	385	15	AAR91442	Human leu8 antigen
41	719	98.1	385	19	AAW60452	Human leu8 antigen
42	719	98.1	385	20	AAW66199	Human leu8 antigen
43	719	98.1	385	21	AAI96138	Human T-cell speci
44	719	98.1	405	22	AAU02447	Human T-lymphocyte
45	718.5	98.0	370	17	AAR98130	Human lymphocyte c

XX	PN	USS514582-A.
XX	FT	Domain
XX	FT	/label= Transmembrane domain. 353..369 /label= Cytoplasmic domain.
XX	FT	/label= Complement binding repeat 2. 330..352
XX	FT	/label= Complement binding repeat 1. 256..314
XX	FT	Binding-site
XX	FT	194..255 /label= EGF domain.
XX	FT	Domain
XX	FT	157..190 /label= Lectin domain.
XX	FT	38..152 /label= Signal region.
XX	FT	Region
XX	FT	1..37 Location/Qualifiers
XX	OS	Homo sapiens.
XX	KX	metastasis.
XX	KW	immunomodulator; cell adhesion; graft rejection; inflammation.
XX	KW	diagnosis; therapy; drug delivery; antiviral; neuroendocrine;
XX	KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX	DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX	DT	01-NOV-1996 (first entry)
XX	AC	AAR98125;
XX	ID	AAR98125 standard; Protein; 369 AA.
XX	RESULT	1

```

XX 07-MAY-1996.
PD
XX 23-FEB-1989; 89US-0315015.
PF
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185670.
XX
PA (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
XX facilitates recovery, improves aqueous solubility and removes
XX potentially immunogenic epitopes. Variants of the human lymphocyte
XX cell surface glycoprotein described in AAR98106 are given in
XX AAR98109-R98135. This variant is a Gly96-Ile97 deletion mutant.
XX
SQ Sequence 369 AA;

Query Match 100.0%; Score 733; DB 17; Length 369;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX immunomodulator; cell adhesion; graft rejection; inflammation;
XX metastasis.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 1..37
XX Domain /label= Signal region.
XX Domain 38..153
XX Domain /label= Lectin domain.
XX Domain 158..191
XX Binding-site /label= EGF domain.
XX Binding-site 195..256
XX Binding-site /label= Complement binding repeat 1.
XX Binding-site 257..315
XX Domain /label= Complement binding repeat 2.
XX Domain 331..353
XX Domain /label= Transmembrane domain.
XX Domain 354..370
XX Domain /label= Cytoplasmic domain.
XX
XX US5514582-A.
XX
XX 07-MAY-1996.
XX
XX 23-FEB-1989; 89US-0315015.
XX
XX 22-NOV-1989; 89US-0440625.
XX 23-FEB-1989; 89US-0315015.
XX 16-DEC-1991; 91US-0808122.
XX 08-DEC-1992; 92US-0986931.
XX 21-JAN-1994; 94US-0185670.
XX
XX (GETH ) GENENTECH INC.
XX
XX Capon DJ, Lasky LA;
XX WPI; 1996-238773/24.
XX
XX Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX binding site of a receptor fused to Ig constant region - useful for
XX diagnosis and treatment e.g. of inflammation
XX
XX Disclosure; Page 19; 41pp; English.
XX
XX A hybrid immunoglobulin chain comprising the ligand binding site of
XX a single transmembrane receptor without an active transmembrane
XX region; fused at its C-terminus with the N-terminus of an
XX immunoglobulin constant region. The receptor is not a member of the
XX immunoglobulin super family, nor a multiple subunit polypeptide
XX encoded by discrete genes. The hybrid immunoglobulin chain combines
XX the adhesion/targeting of a ligand binding partner (LBP) with the
XX effector functions of immunoglobulin and can bind to and/or activate
XX more than one ligand. It can be used diagnostically for the in
XX vitro assay of LBP and their targets; or therapeutically to deliver
XX LBP such as toxins, enzymes, growth factors to particular cells.
XX Typical applications are as antiviral, neuromodulating and
XX immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX The immunoglobulin component increases plasma half life and
XX facilitates purification while deletion of the transmembrane region
XX facilitates recovery, improves aqueous solubility and removes
XX potentially immunogenic epitopes. Variants of the human lymphocyte
XX cell surface glycoprotein described in AAR98106 are given in
XX AAR98109-R98135. This variant is an Asn136 deletion mutant.
XX
SQ Sequence 370 AA;

Query Match 100.0%; Score 733; DB 17; Length 370;

```

Human lymphocyte cell surface glycoprotein (HuLHR) variant.



	Best Local Similarity	100.0%	Pred. No. 1,4e-63;			
Oy	1	CEPTEAPELGTMDCTHPGNGFSFSSQCAFCSCSEGNLTGIEETCGPGFGNWSPEPTCOY	60			
Dd	195	ceplapeelgtmdctchpfngfnfafsqscafcscsegnltgietctcpgfgnwaspptcoy	254			
Oy	61	IQCCEPLSAPDGLIMNCSPHPLASFTSTNACTFTCSGGTELIGKKRTICSSGIMSNPSPIC	120			
Dd	255	iqcceplsapdglimncshplastsftstactticsegteiligkkrticessgiwnpspic	314			
Oy	121	OKLDKSFMIKE	132			
Dd	315	qklDKsfmike	326			
 RESULT 3 AAR98127 standard; Protein: 370 AA.						
XX	AAR98127:					
AC						
DT	01-NOV-1996	(first entry)				
XX						
DE	Human lymphocyte cell surface glycoprotein (HuLHR) variant.					
XX						
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;					
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;					
KW	immunomodulator; cell adhesion; graft rejection; inflammation;					
KW	metastasis.					
OS	Homo sapiens.					
XX						
FH	Key	Location/Qualifiers				
FT	Region	1..37				
FT	Domain	/label= Signal region.				
FT	Domain	38..154				
FT	Domain	/label= Lectin domain.				
FT	Domain	159..191				
FT	Binding-site	/label= EGF domain.				
FT	Binding-site	195..256				
FT	Blinding-site	/label= Complement binding repeat 1.				
FT	Domain	257..315				
FT	Domain	/label= Complement binding repeat 2.				
FT	Domain	331..353				
FT	Domain	/label= Transmembrane domain.				
FT	Domain	354..370				
FT	Domain	/label= Cytoplasmic domain.				
XX						
PN	US5514582-A.					
XX						
PD	07-MAY-1996.					
XX						
PF	23-FEB-1989;	89US-0315015.				
XX						
PR	22-NOV-1989;	89US-0440625.				
PR	23-FEB-1989;	89US-0315015.				
PR	16-DEC-1981;	91US-0808122.				
PR	08-DEC-1982;	92US-0986931.				
PR	21-JAN-1994;	94US-0185670.				
PA	(GETH ) GENENTECH INC.					
PI	Capon DJ, Lasky LA;					
DR	WPI; 1996-238773/24.					
XX						
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand					
PT	binding site of a receptor fused to Ig constant region - useful for					
PT	diagnosis and treatment e.g. of inflammation					
PS	Disclosure; Page 19; 41pp; English.					
XX						

CC	A hybrid immunoglobulin chain comprising the ligand binding site of
CC	a single transmembrane receptor without an active transmembrane
CC	region, fused at its C-terminus with the N-terminus of an
CC	immunoglobulin constant region. The receptor is not a member of the
CC	immunoglobulin super family, nor a multiple subunit polypeptide
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines
CC	the adhesion/targeting of a ligand binding partner (LBP) with the
CC	effector functions of immunoglobulin and can bind to and/or activate
CC	more than one ligand. It can be used diagnostically for the in
CC	vitro assay of LBP and their targets; or therapeutically to deliver
CC	LBP such as toxins, enzymes, growth factors to particular cells.
CC	Typical applications are as antiviral, neuromodulating and
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC	The immunoglobulin component increases plasma half life and
CC	facilitates purification while deletion of the transmembrane region
CC	facilitates recovery, improves aqueous solubility and removes
CC	potentially immunogenic epitopes. Variants of the human lymphocyte
CC	cell surface glycoprotein described in AAR98106 are given in
CC	AAR98109-R98135. This variant is a Ser166 deletion mutant.
CC	
SQ	Sequence 370 AA;
XX	
Query Match	100.0%; Score 733; DB 17; Length 370;
Best Local Similarity	100.0%; Pred. No. 1,4e-63;
Matches 132; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
QY	1 CEPLAPELGTMDCHPFGNFSFSSOCATSCSEGNLNGIEETCGPFGNMSPEPTQY 60
DB	195 CEPLAPELGTMDCHPFGNFSFSSOCATSCSEGNLNGIEETCGPFGNMSPEPTQY 254
QY	61 IQCEPLAPADLIMMCSHPLASFSTACTFTCSSEGTELIGKRTICESSGIMSNPSPIC 120
DB	255 IQCEPLAPDGIIMCSHPLASFSTACTFTCSSEGTELIGKRTICESSGIMSNPSPIC 314
QY	121 QKDKSFSMIKE 132
DB	315 QKDKSFSMIKE 326
RESULT 4	
AAR98109	
ID	AAR98109 standard; Protein; 371 AA.
XX	
AC	AAR98109;
XX	
DT	31-OCT-1996 (first entry)
XX	
DE	Human lymphocyte cell surface glycoprotein (HLHR) variant.
XX	
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW	immunomodulator; cell adhesion; graft rejection; inflammation;
KW	metastasis.
XX	
OS	Homo sapiens.
XX	
FH	Key
FT	Region
FT	/label- Signal region.
FT	38..154
FT	Domain
FT	/label- Lectin domain.
FT	159..192
FT	Domain
FT	/label- EGF domain.
FT	196..257
FT	Binding-site
FT	/label- Complement binding repeat 1.
FT	258..316
FT	Binding-site
FT	/label- Complement binding repeat 2.
FT	332..354
FT	Domain
FT	/label- Transmembrane domain.
FT	355..371
FT	Domain
FT	/label- Cytoplasmic domain.

```

XX XX US514582-A.
XX XX
XX PD 07-MAY-1996.
XX XX
XX PF 23-FEB-1989; 89US-0315015.
XX XX
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX PA
XX (GETH ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX XX
XX DR WPI, 1996-238773/24.
XX XX
XX PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX PT binding site of a receptor fused to Ig constant region - useful for
XX PT diagnosis and treatment e.g. of inflammation
XX PS
XX PS Disclosure; Page 19; 41pp; English.
XX XX
XX CC A hybrid immunoglobulin chain comprising the ligand binding site of
XX CC a single transmembrane receptor without an active transmembrane
XX CC region; fused at its C-terminus with the N-terminus of an
XX CC immunoglobulin constant region. The receptor is not a member of the
XX CC immunoglobulin super family, nor a multiple subunit polypeptide
XX CC encoded by discrete genes. The hybrid immunoglobulin chain combines
XX CC the adhesion/targeting of a ligand binding partner (LBP) with the
XX CC effector functions of immunoglobulin and can bind to and/or activate
XX CC more than one ligand. It can be used diagnostically for the in
XX CC vitro assay of LBP and their targets; or therapeutically to deliver
XX CC LBP such as toxins, enzymes, growth factors to particular cells.
XX CC Typical applications are as antiviral, neuromodulating and
XX CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX CC The immunoglobulin component increases plasma half life and
XX CC facilitates purification while deletion of the transmembrane region
XX CC facilitates recovery, improves aqueous solubility and removes
XX CC potentially immunogenic epitopes. Variants of the human lymphocyte
XX CC cell surface glycoprotein described in AAR98106 are given in
XX CC AAR98109-R98135. This variant contains a Arg98-Asp59; Lys-Glu
XX CC substitution.
XX XX
XX SQ Sequence 371 AA;

```

```

DT 01-NOV-1996 (first entry)
XX XX
XX DE Human lymphocyte cell surface glycoprotein (HLHR) variant.
XX XX
XX KW Immunoglobulin; transmembrane receptor; adhesion; targeting;
XX KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;
XX KW immunomodulator; cell adhesion; graft rejection; inflammation;
XX KW metastasis.
XX OS Homo sapiens.
XX XX
XX FH Key
XX FH Region
XX FH
XX FH Domain
XX FH
XX FH Domain
XX FH
XX FH Binding-site
XX FH
XX FH Binding-site
XX FH
XX FH Binding-site
XX FH
XX FH Domain
XX FH
XX PN US514582-A.
XX XX
XX PD 07-MAY-1996.
XX XX
XX PF 23-FEB-1989; 89US-0315015.
XX XX
XX PR 22-NOV-1989; 89US-0440625.
XX PR 23-FEB-1989; 89US-0315015.
XX PR 16-DEC-1991; 91US-0808122.
XX PR 08-DEC-1992; 92US-0986931.
XX PR 21-JAN-1994; 94US-0185670.
XX PA
XX (GETH ) GENENTECH INC.
XX PI Capon DJ, Lasky LA;
XX XX
XX DR WPI, 1996-238773/24.
XX XX
XX PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX PT binding site of a receptor fused to Ig constant region - useful for
XX PT diagnosis and treatment e.g. of inflammation
XX PS
XX PS Disclosure; Page 19; 41pp; English.
XX XX
XX CC A hybrid immunoglobulin chain comprising the ligand binding site of
XX CC a single transmembrane receptor without an active transmembrane
XX CC region; fused at its C-terminus with the N-terminus of an
XX CC immunoglobulin constant region. The receptor is not a member of the
XX CC immunoglobulin super family, nor a multiple subunit polypeptide
XX CC encoded by discrete genes. The hybrid immunoglobulin chain combines
XX CC the adhesion/targeting of a ligand binding partner (LBP) with the
XX CC effector functions of immunoglobulin and can bind to and/or activate
XX CC more than one ligand. It can be used diagnostically for the in
XX CC vitro assay of LBP and their targets; or therapeutically to deliver
XX CC LBP such as toxins, enzymes, growth factors to particular cells.
XX CC Typical applications are as antiviral, neuromodulating and
XX CC immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX CC treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX CC The immunoglobulin component increases plasma half life and
XX CC facilitates purification while deletion of the transmembrane region
XX CC facilitates recovery, improves aqueous solubility and removes
XX CC potentially immunogenic epitopes. Variants of the human lymphocyte
XX CC cell surface glycoprotein described in AAR98106 are given in
XX CC AAR98109-R98135. This variant contains an Ala71Ser substitution.
XX XX
XX SQ Sequence 371 AA;

```

Query match	Score	DB	Length
100.0%	733	17	371

Best Local Similarity 100.0%; Pred. No. 1.4e-03;  
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 CEPLEAELGMOCTHFGNPNSSFGCAFQSGSEBTLNLTGIEHTTCGPRGMMSSPEPCQY 60
        |||||
        196 CEPLEAELGUMDCDTHPGNIFSGQCAFQSGSEBTLNLTGIEHTTCGPRGMMSSPEPCQY 255
        |||||
QY      61 IQCEPLAPDLGINCNCHPLASFSTACMFICSEGTLELGGKKTKTIOESSGIMSNBPCIC 120
        |||||
        256 IQCEPLAPDLGINCNCHPLASFSTACMFICSEGTLELGGKKTKTIOESSGIMSNBPCIC 315
        |||||
QY      121 QKLDKSFMSIKE 132
        |||||
Db      316 qkldksfsmike 327
```

## RESULT 6

AA98111  
ID AA98111 standard; Protein; 371 AA.

AC AAR98111;

DT 01-NOV-1996 (first entry)

Human lymphocyte surface glycoprotein (HLHR) variant.

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
KW immunomodulator; cell adhesion; graft rejection; inflammation  
KW metastasis.

OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2
FT	Domain	337..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.

PN US5514582-A.

PD 07-MAY-1996.

PF 23-FEB-1989; 89US-0315015.

PR 22-NOV-1989; 89US-0440625.

PR 16-DEC-1991; 91US-0808122.

PR 21-JAN-1994; 94US-0185670.

PA (GETH ) GENENTECH INC.

PI Capon DJ, Lasky LA;

DR WPI; 1996-238773/24.

PT Nucleic acid encoding

PT diagnosis and treatment e.g. of inflammation

XX	Disclosure; Page 19; 41pp; English
PS	

A hybrid immunoglobulin chain comprising the ligand binding site of a single transmembrane receptor without an active transmembrane region, fused at its C-terminus with the N-terminus of an immunoglobulin constant region. The receptor is not a member of the immunoglobulin super family, nor a multiple subunit polypeptide encoded by discrete genes. The hybrid immunoglobulin chain combines the adhesion/targeting of a ligand binding partner (LBP) with the effector functions of immunoglobulin and can bind to and/or activate more than one ligand. It can be used diagnostically for the *in vitro* assay of LBP and their targets, or therapeutically to deliver LBP such as toxins, enzymes, growth factors to particular cells. Typical applications are as antiviral, neuromodulating and immunomodulating agents, or as modulators of cell adhesion (e.g. in treating graft rejection, inflammation, metastasis of lymphoma etc.) The immunoglobulin component increases plasma half life and facilitates purification while deletion of the transmembrane region facilitates recovery, improves aqueous solubility and removes potentially immunogenic epitopes. Variants of the human lymphocyte cell surface glycoprotein described in AAR8106 are given in AAR98109-998135. This variant contains an Lys78Gln substitution.

**SQ Sequence 371 AA;**

Query Match	100.0%	Score 733;	DB 17;	Length 371;
Best Local Similarity	100.0%	Pred. No. 1.4e-63;		
Matches 132; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

QY 1 CEPLEAPRLGIMDCDTHPRGNFSFSSQCAFSCSSESTNLGTLEETTCGFFGMMSSPEPICQY 60

Dp cepleaprlgimdcDthprgnfsfssqcafscssegtnltgtleetcsgpfgmmsspepicqy 255

QY 61 IQCEPLSAPDLGIMNCSHPLASFSTFACPTICSEBGEHLGKKTTCOESSGIMSNBPC 120

Dp iqceplsapdlgimncshplasfstfacticsegehlgkkticoessgimnsbpc 315

QY 121 OKLKSFSMIKE 132

Dp oklksfstmike 327

## RESULT 7

AA98112  
ID AA98112 standard; Protein; 371 AA.

AC AAR98112;

DT 01-NOV-1996 (first entry)

DE Human lymphocyte cell surface glycoprotein (HULHR) variant.

KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.

OS Homo sapiens.

FT	Key	Location/Qualifiers
FT	Region	1..37
FT	/label=	Signal region.
FT	Domain	38..154
FT	/label=	Lectin domain.
FT	Domain	159..192
FT	/label=	ECF domain.
FT	Binding-site	196..257
FT	/label=	Complement binding repeat 1
FT	Binding-site	256..316
FT	/label=	Complement binding repeat 2
FT	Domain	332..354

FT		/label= Transmembrane domain.
FT	Domain	355..371
PX		/label= Cytoplasmic domain.
PN		
XX	US5514582.A.	
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
XX	21-JAN-1994;	94US-0185670.
PA	(GETH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI; 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
PS	Disclosure: Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-R98135. This variant contains an Asp116Glu substitution.	
XX		
SQ	Sequence 371 AA:	
	Query Match 100.0%; Score 733; DB 17; Length 371;	
	Best Local Similarity 100.0%; Pred. No. 1,4e-63;	
	Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps	
QY	1 CEPLAEPLGTMDCTPFGNFSFSSOCACSGEQTULGIETFCQFENMSSPEPTCY 60	
DB	196 ceplaeplgundctlpfnfsfscacscsegtullyeeetcpqpnwsspeptcy 255	
QY	61 IOCEPLASAPDLGIMNCSHPLASFSPFSACTFFICSEGETELIGKKRTICSSGGIWSNSPIC 120	
DB	256 iocleplsapdlgimncshplasfstfsactfficsegteiligkktlccessgglwnspic 315	
QY	121 QKIDKSFSMIKE 132	
DB	316 qkidksfsmike 327	
	RESULT 8	
	AAR98113	
	ID AAR98113 standard; Protein; 371 AA.	
	XX	

AC	AA98113;	
DT	01-NOV-1996	(first entry)
XX	Human lymphocyte cell surface glycoprotein (HULHR) variant.	
XX		
DE	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KW	immunomodulator; cell adhesion; graft rejection; inflammation;	
KW	metastasis.	
OS	Homo sapiens.	
XX		
XX		
FH	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= BGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
XX		/label= Cytoplasmic domain.
PN	US514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PF	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0801812.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
PA	(GETH ) GENENTECH INC.	
XX		
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI: 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
XX		
XX	Disclosure; Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AA98106 are given in	
CC	AA98109-R98135. This variant contains a Leu150Val substitution.	

xx	Sequence	371 AA:
xx	Query Match	100.0%; Score 733; DB 17; Length 371;
xx	Best Local Similarity	100.0%; Pred. No. 1,4e-63;
xx	Matches 132; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 CEPLEAPELGIMDCIHPNENFSFSSQCAFSCSEGTNLGIIEFTTCGPFGNWSSPEPTCGV	60
Db	196 cepleapelgimdcuhptcfnfgfssqcafcscsegtnlgtleetcpgfgnwaspeptcgv	255
OY	61 IOCEPLAAPPDGLMNCNSHFLASFEFTSACTFTCSGEETLIGKKKTTICCSSGITMSNPSPIC	120
Db	256 lqceplsapdylgimcnshflasftstactficsegtellgkkticcssgitmsnpspic	315
OY	121 OKDKSFWSMIKE 132	
Db	316 qklidsfsmike 327	
RESULT 9	AAR98114	
ID	AAR98114 standard; Protein; 371 AA.	
XX AC	AAR98114;	
XX DT	01-NOV-1996 (first entry)	
DE	Human lymphocyte cell surface glycoprotein (HLuHR) variant.	
XX		
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;	
RW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KW	immunomodulator; cell adhesion; graft rejection; inflammation;	
XX	metastasis.	
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..37
FT	Domain	/label= Signal region.
FT	Domain	38..154
FT	Domain	/label= Lectin domain.
FT	Binding-site	159..192
FT	Binding-site	/label= EGF domain.
FT	Binding-site	196..257
FT	Binding-site	/label= Complement binding repeat 1.
FT	Domain	258..316
FT	Domain	/label= Complement binding repeat 2.
FT	Domain	332..354
FT	Domain	/label= Transmembrane domain.
FT	Domain	355..371
FT	Domain	/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PE	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI: 1996-238773/24.	
XX	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	

Query Match	100.0%	Score 733;	DB 17;	Length 371;
Best Local Similarity	100.0%	Pre. No. 1.4e-63;		
Matches 132;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY 1	CEPLEAPELGTMDCTPHFNGFNFSFSSQCAPSCSGTNLTGIEFTGCPFGNMSSPREPTCOV	60		
DB 196	cepleapeelgtmdctphfngfnfsfsgqarscsegtnltgieetcgpfgnwspeptcqv	255		
QY 61	IOCEPLAPDLGIMNCSHPLASFSFPTSACTFTICSEGTETLIGKRRKTCSSGIMSNPSPIC	120		
DB 256	iqceplsapdldgimncshplasfsfptaactfticsegtetligrkkticessgilmnsnpic	315		
QY 121	QKLDKSFMSMIKE 132			
DB 316	qkldksfmsmike 327			
RESURF_10				
AAK98115				
ID	AAK98115 standard; Protein; 371 AA.			
XX				
AC	AAK98115;			
XX				
DT	01-NOV-1996 (first entry)			
XX				
DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.			
XX				
KW	Immunoglobulin; transmembrane receptor; adhesion; targeting;			
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;			
KW	immunomodulator; cell adhesion; graft rejection; inflammation;			
KM	metastasis.			
XX				
OS	Homo sapiens.			
XX				
FT	Key	Location/Qualifiers		
FH	Region	/label= Signal region.		
FT	Domain	38..154		
FT	Domain	/label= Lectin domain.		
FT	Domain	159..192		
FT	Binding-site	/label= EGF domain.		
FT	Binding-site	196..257		
FT	Binding-site	/label= Complement binding repeat 1.		
FT	Binding-site	258..316		

```

FT      Domain                               /label= Complement binding repeat 2.
FT      .332..354
FT      /label= Transmembrane domain.
FT      Domain                               /label= Cytoplasmic domain.
FT      .355..371
XX      US5514582-A.
XX      07-MAY-1996.
XX      23-FEB-1989; 89US-0315015.
XX      22-NOV-1989; 89US-0440625.
XX      23-FEB-1989; 89US-0315015.
XX      16-DEC-1991; 91US-0808122.
XX      08-DEC-1992; 92US-0986931.
XX      21-JAN-1994; 94US-0185670.
XX      (GETH ) GENENTECH INC.
XX      Capon DJ, Lasky LA;
XX      WPI; 1996-238773/24.
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX      Disclosure; Page 19; 41pp; English.
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targeting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate
CC      more than one ligand. It can be used diagnostically for the in
CC      vitro assay of LBP and their targets; or therapeutically to deliver
CC      LBP such as toxins, enzymes, growth factors to particular cells.
CC      Typical applications are as antiviral, neuromodulating and
CC      immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC      The immunoglobulin component increases plasma half life and
CC      facilitates purification while deletion of the transmembrane region
CC      facilitates recovery, improves aqueous solubility and removes
CC      potentially immunogenic epitopes. Variants of the human lymphocyte
CC      cell surface glycoprotein described in AAR98106 are given in
CC      AAR98109-R98135. This variant contains an Ile174Ileu substitution.
XX      Sequence ` 371 AA;
SQ

```

```

ID      AAR98116 standard; Protein; 371 AA.
XX      AAR98116;
AC      01-NOV-1996 (first entry)
XX      Human lymphocyte cell surface glycoprotein (HULHR) variant.
XX      DE
XX      KW      Immunoglobulin; transmembrane receptor; adhesion; targeting;
KW      diagnosis; therapy; drug delivery; antiviral; neuromodulator;
KW      immunomodulator; cell adhesion; graft rejection; inflammation;
KW      metastasis.
XX      Homo sapiens.
XX      OS
XX      FH      Key
XX      FT      Region
XX      FT      1..37
XX      FT      /label= Signal region.
XX      FT      38..154
XX      FT      /label= Lectin domain.
XX      FT      159..192
XX      FT      /label= EGF domain.
XX      FT      196..257
XX      FT      /label= Complement binding repeat 1.
XX      FT      258..316
XX      FT      /label= Complement binding repeat 2.
XX      FT      332..354
XX      FT      /label= Transmembrane domain.
XX      FT      355..371
XX      FT      /label= Cytoplasmic domain.
XX      US5514582-A.
XX      07-MAY-1996.
XX      23-FEB-1989; 89US-0315015.
XX      22-NOV-1989; 89US-0440625.
XX      23-FEB-1989; 89US-0315015.
XX      16-DEC-1991; 91US-0808122.
XX      08-DEC-1992; 92US-0986931.
XX      21-JAN-1994; 94US-0185670.
XX      (GETH ) GENENTECH INC.
XX      Capon DJ, Lasky LA;
XX      WPI; 1996-238773/24.
XX      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
PT      binding site of a receptor fused to Ig constant region - useful for
PT      diagnosis and treatment e.g. of inflammation
XX      Disclosure; Page 19; 41pp; English.
XX      A hybrid immunoglobulin chain comprising the ligand binding site of
CC      a single transmembrane receptor without an active transmembrane
CC      region; fused at its C-terminus with the N-terminus of an
CC      immunoglobulin constant region. The receptor is not a member of the
CC      immunoglobulin super family, nor a multiple subunit polypeptide
CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
CC      the adhesion/targeting of a ligand binding partner (LBP) with the
CC      effector functions of immunoglobulin and can bind to and/or activate
CC      more than one ligand. It can be used diagnostically for the in
CC      vitro assay of LBP and their targets; or therapeutically to deliver
CC      LBP such as toxins, enzymes, growth factors to particular cells.
CC      Typical applications are as antiviral, neuromodulating and
CC      immunomodulating agents, or as modulators of cell adhesion (e.g. in
CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
CC      The immunoglobulin component increases plasma half life and
CC      facilitates purification while deletion of the transmembrane region
CC      facilitates recovery, improves aqueous solubility and removes
CC      potentially immunogenic epitopes. Variants of the human lymphocyte

```

RESULT 11  
AAR98116



FT	Key	Location/Qualifiers
FT	Peptide	1..38
FT		/label= signal_sequence
FT	Protein	39..372
FT		/note="trp39 is probable N-terminus of mature LHR"
FT	Domain	39..135
FT		/label= Lectin_domain
FT	Modified-site	60..62
FT		/note="potential N-glycosylation site"
FT	Modified-site	104..106
FT		/note="potential N-glycosylation site"
FT	Domain	160..193
FT		/label= BGF_domain
FT	Modified-site	177..179
FT		/note="potential N-glycosylation site"
FT	Region	197..258
FT		/label= Complement_Binding_Repeat_1
FT	Modified-site	216..218
FT		/note="potential N-glycosylation site"
FT	Modified-site	232..234
FT		/note="potential N-glycosylation site"

Query Match	100.0%	Score 733:	DB 14:	Length 372:
Best Local Similarity	100.0%	Pred. No. 1,4e-63:		
Matches 132:	Conservative	0:	Mismatches 0:	Indels 0:
				Gaps 0:
QY . 1	CEPLEAPELGTMDCHHPFEGNFSSQCAFCSCSEGNINGTEETTCGPFGNMSSPEPTCOV	60		
Db 197	cepleapeelgmdcthrpfnfissqcatcscgegnlgyteetctgpfngwsspeptcyv	256		
QY 61	IQCEPLSAPDLGIMNCSHPLASFTSACTFICSEGTETIGKKKTIICSSGIMSNPSPIC	120		
Db 257	iqceplsapdldgimncshplasftsactficsegtetlqgkkkticessgimwnspic	316		
QY 121	QKLDKSPFSMIKE 132			
Db 317	qkldksfsmike 328			

RESULT	15
AAR38908	
ID	AAR38908 standard; Protein; 372 AA
XX	



AC AAR38908;  
 XX  
 DT 11-JAN-1994 (first entry)  
 XX  
 DE HuLHR.  
 XX  
 KW Human; murine; lymphocyte; cell surface glycoprotein; homing receptor;  
 KW LHR; endothelium; lymphoid tissue; signal; domain; complement binding;  
 KW carbohydrate binding; epidermal growth factor-like; egf; intracellular;  
 KW transmembrane binding; cytoplasmic; ligand binding partner protein;  
 KW TMD; LBPP.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide 20..32 Location/Qualifiers  
 FT /note="Signal peptide"  
 FT Protein 39..372  
 FT /note="Mature protein"  
 FT Domain 39..155  
 FT /note="Lectin domain"  
 FT Domain 160..193  
 FT /note="egf domain"  
 FT Domain 197..317  
 FT /note="Complement factor binding domain"  
 FT Domain 333..355  
 FT /note="Transmembrane binding domain"  
 FT Domain 356..372  
 FT /note="Cytoplasmic domain"  
 XX  
 PN US522538-A.  
 XX  
 PD 06-JUL-1993.  
 XX  
 PF 23-FEB-1989; 8905-0315015.  
 XX  
 PR 23-FEB-1989; 8905-0315015.  
 PR 22-NOV-1989; 8905-0440625.  
 PR 16-DEC-1991; 9105-0808122.  
 XX  
 PA (GETH ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI: 1993-226664/28.  
 DR N-PSDB; AAQ44243.  
 XX  
 PT New lymphocyte homing receptor immunoglobulin fusion  
 PT polypeptide(s) - used to inhibit binding of lymphocytes in  
 PT therapeutic and diagnostic uses  
 XX  
 PS Disclosure; Fig 1; 44pp; English.  
 XX  
 XX The sequences given in AAR38908-09 represent human and murine lymphocyte  
 CC cell surface glycoprotein (LHR) respectively. These proteins mediate  
 CC the binding of lymphocytes to the endothelium of lymphoid tissue. LHR  
 CC is a glycoprotein which contains a signal domain, a carbohydrate  
 CC binding domain, an epidermal growth factor-like (egf) domain, at least  
 CC one complement binding domain repeat, a transmembrane binding domain  
 CC (TMD) and a charged intracellular or cytoplasmic domain. The murine  
 CC and human amino acid sequences show a high degree of overall homology  
 CC (83%), however degrees of homology between the various domains is  
 CC variable. These proteins may be used to a ligand binding partner  
 CC protein (LBPP) which causes an increase in the half life of the LHR.  
 CC The fusions may be used therapeutically to compete with the normal  
 CC binding of lymphocytes to lymphoid tissue. They may be used in organ  
 CC or graft rejection and for the treatment of inflammation.  
 XX  
 SQ Sequence 372 AA;

Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CEPLAPRLGTMDCITPRGNSFSQCAFSCSSEGNLTGIRETTGCPRGMSNPPTQY 60  
 DB 197 cepleaprlgimdcitprgnfsqcafsccsegnltgirettcpgrgmsnpptqv 256  
 QY 61 IQCEPLSPDGLIMNCSPHPLASFSTSACTFICSGTELGKKTICSSGIMWSPIC 120  
 DB 257 iqceplspdglimncsphplasfstsaactficsgtelegkkticssgimwspic 316  
 QY 121 QKLDKSFMSIKE 132  
 DB 317 qkldksfmsike 328  
 RESULT 16  
 AAR76506  
 ID AAR76506 standard; Protein: 372 AA.  
 XX  
 AC AAR76506;  
 XX  
 DT 01-DEC-1995 (first entry)  
 XX  
 DE Human LHR.  
 XX  
 KW Lymphocyte homing receptor; lymphocyte cell surface glycoprotein;  
 KW LHR; ligand binding partner; immunoglobulin; constant region;  
 KW antibody engineering; immunomodulator.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FT Peptide 20..32 Location/Qualifiers  
 FT /label="Sig-peptide"  
 FT /note="hydrophobic domain, may act as signal  
 FT for insertion into the endoplasmic  
 FT reticulum lumen"  
 FT Protein 39..372  
 FT /label="Mat-protein"  
 FT Domain 39..334  
 FT /note="putative extracellular domain"  
 FT Modified-site 60..62  
 FT /label="N-glycosylation-site"  
 FT Modified-site 104..106  
 FT /label="N-glycosylation-site"  
 FT Modified-site 177..179  
 FT /label="N-glycosylation-site"  
 FT Modified-site 216..218  
 FT /label="N-glycosylation-site"  
 FT Modified-site 232..234  
 FT /label="N-glycosylation-site"  
 FT Modified-site 246..248  
 FT /label="N-glycosylation-site"  
 FT Modified-site 271..273  
 FT /label="N-glycosylation-site"  
 FT Modified-site 311..313  
 FT /label="N-glycosylation-site"  
 FT Domain 335..357  
 FT /note="putative stop transfer or membrane anchor  
 FT domain"  
 FT Region 358..372  
 FT /note="putative intracellular region"  
 XX  
 PN US5428130-A.  
 XX  
 PD 27-JUN-1995.  
 XX  
 PE 23-FEB-1989; 8905-0315015.  
 XX  
 PR 22-NOV-1989; 8905-0440625.  
 PR 23-FEB-1989; 8905-0315015.  
 PR 16-DEC-1991; 9105-0808122.  
 PR 08-DEC-1992; 9205-0986931.

```

XX (GETH ) GENENTECH INC.
PA
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI: 1995-240086/31.
DR
XX N-PSDB; AAQ92802.
XX
XX New hybrid ligand binding partner molecules - fused to immunoglobulin
PT constant region sequences to increase stability and in vivo plasma
PT half-life
XX
XX Disclosure; Fig.1a-1c; 40pp; English.
XX
XX A murine Mel 14 antigen cDNA clone was used to screen a lambda gt10
CC cDNA library derived from human peripheral blood lymphocyte mRNA
CC obtd. from primary cells. A cDNA clone encoding LHR was isolated.
XX
SQ Sequence 372 AA;

Query Match 100.0%; Score 733; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPDLGMDCTHPFGNPSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
DB 197 ceplapdlgmdcthpfgnfisfsqcafscegtnlgieettcgpgnmspeptcqv 256
OY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIWNSPSPIC 120
DB 257 igceplapdlgimncshplasfstsactficsegtteigkkticssgiwnspspic 316
OY 121 QKDKSFSMIKE 132
DB 317 qkdkfsfsmike 328

RESULT 17
AAR83050
ID AAR83050 standard; Protein; 372 AA.
XX
XX AAR83050;
AC
XX
XX 31-JAN-1996 (first entry)
DT
XX
XX Human LHR.
DE
XX
XX Lymphocyte cell surface glycoprotein; LHR; transmembrane receptor;
KW immunoglobulin; IgG; constant region; receptor-mediated disease;
KW vector; plasma-life.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
FH 20..32
FH Peptide /label= Sig-peptide
FT 39..155
FT Domain /label= Carbohydrate_binding_domain
FT 160..193
FT Domain /label= Epidermal_growth_factor_domain
FT 197..317
FT Domain /label= Complement_factor_binding_domain
FT 333..355
FT Domain /label= Transmembrane_binding_domain
FT 356..372
FT Domain /label= Cytoplasmic_domain
FT /note= "potential stop transfer sequence"
FT Modified-site 60..62
FT /label= N-glycosylation_site
FT Modified-site 104..106
FT /label= N-glycosylation_site
FT Modified-site 177..179

```

```

FT /label= N-glycosylation_site
FT 216..218
FT Modified-site /label= N-glycosylation_site
FT 232..234
FT Modified-site /label= N-glycosylation_site
FT 248..248
FT Modified-site /label= N-glycosylation_site
FT 271..273
FT Modified-site /label= N-glycosylation_site
FT 311..313
FT /label= N-glycosylation_site

US5455165-A.
PN
XX
XX 03-OCT-1995.
PD
XX
XX 23-FEB-1989; 89US-0315015.
PF
XX
XX 22-NOV-1989; 89US-0440625.
PR 23-FEB-1989; 89US-0315015.
PR 16-DEC-1991; 91US-0808122.
PR 08-DEC-1992; 92US-0986931.
PR 21-JAN-1994; 94US-0185669.
XX
XX (GETH ) GENENTECH INC.
PA
XX
XX Capon DJ, Lasky LA;
PI
XX
XX WPI: 1995-350776/45.
DR N-PSDB; AAT05869.
XX
XX Expression vector encoding fusion protein to increase plasma life -
PT comprises receptor ligand binding site and Ig constant region, for
PT treatment of receptor mediated disease
XX
XX Disclosure; Fig 1; 42pp; English.
XX
XX A mouse LHR (lymphocyte cell surface glycoprotein) cDNA clone was used
CC to screen an oligo-dT primed lambda gt10 cDNA library derived from
CC human peripheral blood lymphocyte mRNA obtd. from primary cells. A
CC 2.2 kb clone (sequence given in AAT05869) was isolated that encoded the
CC human LHR protein (AAR83050). LHR-IgG hybrids were constructed for use
CC in the targeting of therapeutic moieties to lymphoid tissue.
XX
XX Sequence 372 AA;
SQ

Query Match 100.0%; Score 733; DB 16; Length 372;
Best Local Similarity 100.0%; Pred. No. 1.4e-63;
Matches 132; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPDLGMDCTHPFGNPSFSSQCAFSCSEGTNLGIEETTCGPGNMSPEPTCOV 60
DB 197 ceplapdlgmdcthpfgnfisfsqcafscegtnlgieettcgpgnmspeptcqv 256
OY 61 IOCEPLAPDLGIMNCSHPLASFSTACTFICSEGTTELIGKKTKICSSGIWNSPSPIC 120
DB 257 igceplapdlgimncshplasfstsactficsegtteigkkticssgiwnspspic 316
OY 121 QKDKSFSMIKE 132
DB 317 qkdkfsfsmike 328

RESULT 18
AAR98106
ID AAR98106 standard; Protein; 372 AA.
XX
XX AAR98106;
AC
XX
XX 31-OCT-1996 (first entry)
DT
XX
XX Human lymphocyte cell surface glycoprotein (HuLHR).
DE

```



WPI; 1999-034122/03  
N-PSDB; AAV08321.

Domain	/label= Transmembrane domain. 358..374
Domain	





```

FT      XX      /label= Cytoplasmic domain.
XX      PN      USS514582-A.
XX      PD      07-MAY-1996.
XX      PF      23-FEB-1989;      89US-0315015.
XX      PR      22-NOV-1989;      89US-0440625.
XX      PR      23-FEB-1989;      89US-0315015.
XX      PR      16-DEC-1991;      91US-0808122.
XX      PR      08-DEC-1992;      92US-0986931.
XX      PR      21-JAN-1994;      94US-0185670.
XX      PA      (GETH ) GENENTECH INC.
XX      PI      Capon DJ, Lasky LA:
XX      DR      WPI: 1996-238773/24.
XX      PT      Nucleic acid encoding hybrid immunoglobulin comprising the ligand
XX      PT      binding site of a receptor fused to Ig constant region - useful for
XX      PS      diagnosis and treatment e.g. of inflammation
XX      PS      Disclosure: Page 19; 41pp; English.
XX      CC      A hybrid immunoglobulin chain comprising the ligand binding site of
XX      CC      a single transmembrane receptor without an active transmembrane
XX      CC      region, fused at its C-terminus with the N-terminus of an
XX      CC      immunoglobulin constant region. The receptor is not a member of the
XX      CC      immunoglobulin super family, nor a multiple subunit polypeptide
XX      CC      encoded by discrete genes. The hybrid immunoglobulin chain combines
XX      CC      the adhesion/targeting of a ligand binding partner (LBP) with the
XX      CC      effector functions of immunoglobulin and can bind to and/or activate
XX      CC      more than one ligand. It can be used diagnostically for the in
XX      CC      vitro assay of LBP and their targets; or therapeutically to deliver
XX      CC      LBP such as toxins, enzymes, growth factors to particular cells.
XX      CC      Typical applications are as antiviral, neuromodulating and
XX      CC      immunomodulating agents, or as modulators of cell adhesion (e.g. in
XX      CC      treating graft rejection; inflammation; metastasis of lymphoma etc.)
XX      CC      The immunoglobulin component increases plasma half life and
XX      CC      facilitates purification while deletion of the transmembrane region
XX      CC      facilitates recovery, improves aqueous solubility and removes
XX      CC      potentially immunogenic epitopes. Variants of the human lymphocyte
XX      CC      cell surface glycoprotein described in AAR98106 are given in
XX      CC      AAR98109-R98135. This variant contains an Ile302Leu substitution.
XX      SO      Sequence      371 AA:

Query Match      99.7%; Score 731; DB 17; Length 371;
Best Local Similarity 99.2%; Pred. No. 2,2e-63;
Matches 131; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

      1 CEPLEAPETGTMDCTHPFGNFFSFSQCAFSCEGTNLTIGIEETGCPFGFMSSPEPTCOV 60
      196 cepleapelgtmdctchpfnfsgfsgcasccegnlgtieetogfpfnwsspeptcy 255
      61 IOCEPLSAPADLCIMNCSPHSLASFSTPSACTFTCSGCTELIGKKTKTICSSSGIWSNDSPIC 120
      256 lqceplsapdqlimncshp lasfstpsactftcsegte llgkkkticsssg iwsnspic 315
      121 OKLDSFSMIKE 132
      316 qkldsfsmike 327

RESULT 25
AAR98119
ID      AAR98119 standard; Protein; 371 AA.
AC      AAR98119;
XX

```

DT	01-NOV-1996	(first entry)
XX		
DE	Human lymphocyte cell surface glycoprotein (HULHR) variant.	
XX		
KM	Immunoglobulin: transmembrane receptor; adhesion; targeting;	
KM	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KM	immunomodulator; cell adhesion; graft rejection; inflammation;	
XX	metastasis.	
OS		
XX	Homo sapiens.	
PH		
FT	Key	
FT	Location/Qualifiers	
FT	1..37	
FT	/label= Signal region.	
FT	38..154	
FT	/label= Lectin domain.	
FT	159..192	
FT	/label= EGF domain.	
FT	196..257	
FT	/label= Complement binding repeat 1.	
FT	258..316	
FT	/label= Complement binding repeat 2.	
FT	332..354	
FT	/label= Transmembrane domain.	
FT	355..371	
FT	/label= Cytoplasmic domain.	
PN		
XX	US5514582-A.	
XX		
PD	07-MAY-1996.	
PF		
XX	23-FEB-1989;	
XX	89US-0315015.	
PR	22-NOV-1989;	
PR	89US-0440625.	
PR	23-FEB-1989;	
PR	89US-0315015.	
PR	16-DEC-1991;	
PR	91US-0808122.	
PR	08-DEC-1992;	
PR	92US-0986931.	
XX	21-JAN-1994;	
XX	94US-0185670.	
PA		
XX	(GETH ) GENENTECH INC.	
PI		
XX	Capon DJ, Lasky LA;	
DR		
XX	WPI: 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
XX	diagnosis and treatment e.g. of inflammation	
PS		
XX	Disclosure: Page 19; 41pp; English.	
CC		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain comprises	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one LBP. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	creating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-RB8135. This variant contains a Ser226Thr substitution.	
XX		
XX	Sequence 371 AA:	

Query Match 99.6%; Score 730; DB 17; Length 371;  
 Best Local Similarity 99.2%; Pred. No. 2.8e-63;  
 Matches 131; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPELGTMDCIHPGNSFSFSCAFSCSEGTNLGIEPTTCGPGNMSPEPTCOV 60  
 |||  
 DB 196 ceplapeligtmcdchpignfsfsqcafcscgntlgieetccpgtgnwsspccqv 255  
 |||  
 OY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTGCTGCTGLKKTTCSSGSGWSNPSPIC 120  
 |||  
 DB 256 iqceplapdlgimncshplastfctactlcegtellgkktlccssgswnpapic 315  
 |||  
 OY 121 QKDKSSSMIKE 132  
 |||  
 DB 316 qkdkssismike 327  
 |||

RESULT 26  
 ID AAR98117 standard; Protein; 371 AA.  
 AC AAR98117;  
 DT 01-NOV-1996 (first entry)  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..37  
 FT Domain /label= Signal region.  
 FT Domain 38..154  
 FT Domain /label= Lectin domain.  
 FT Domain 159..192  
 FT Binding-site /label= EGF domain.  
 FT Binding-site 196..257  
 FT Binding-site /label= Complement binding repeat 1.  
 FT Domain 258..316  
 FT Domain /label= Complement binding repeat 2.  
 FT Domain 332..354  
 FT Domain /label= Transmembrane domain.  
 FT Domain 355..371  
 FT Domain /label= Cytoplasmic domain.  
 PN US5514582-A.  
 XX  
 PD 07-MAY-1996.  
 XX  
 PF 23-FEB-1989; 89US-0315015.  
 XX  
 PR 22-NOV-1989; 89US-0440625.  
 PR 23-FEB-1989; 89US-0315015.  
 PR 16-DEC-1991; 91US-0808122.  
 PR 08-DEC-1992; 92US-0986931.  
 PR 21-JAN-1994; 94US-0185670.  
 XX  
 PA (GENE ) GENENTECH INC.  
 XX  
 PI Capon DJ, Lasky LA;  
 XX  
 DR WPI; 1996-238773/24.  
 XX  
 PT Nucleic acid encoding hybrid immunoglobulin comprising the ligand  
 PT binding site of a receptor fused to Ig constant region - useful for  
 PT diagnosis and treatment e.g. of inflammation

XX  
 PS Disclosure; Page 19; 41pp; English.  
 CC  
 CC A hybrid immunoglobulin chain comprising the ligand binding site of  
 CC a single transmembrane receptor without an active transmembrane  
 CC region; fused at its C-terminus with the N-terminus of an  
 CC immunoglobulin constant region.  
 CC immunoglobulin super family. The receptor is not a member of the  
 CC the adhesion/targeting of a ligand binding partner (LBP) with the  
 CC effector functions of immunoglobulin and can bind to and/or activate  
 CC more than one ligand. It can be used diagnostically for the in  
 CC vitro assay of LBP and their targets; or therapeutically to deliver  
 CC LBP such as toxins, enzymes, growth factors to particular cells.  
 CC Typical applications are as antiviral, neuromodulating and  
 CC immunomodulating agents, or as modulators of cell adhesion (e.g. in  
 CC the immunoglobulin component increases plasma half life and  
 CC facilitates graft rejection; inflammation; metastasis of lymphoma etc.)  
 CC facilitates recovery, improves aqueous solubility and removes  
 CC potentially immunogenic epitopes. Variants of the human lymphocyte  
 CC cell surface glycoprotein described in AAR98106 are given in  
 CC AAR98109-R98135. This variant contains a Thr211Ser substitution.  
 XX  
 SQ Sequence 371 AA;

Query Match 99.5%; Score 729; DB 17; Length 371;  
 Best Local Similarity 99.2%; Pred. No. 3.5e-63;  
 Matches 131; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CEPLAPELGTMDCIHPGNSFSFSCAFSCSEGTNLGIEPTTCGPGNMSPEPTCOV 60  
 |||  
 DB 196 ceplapeligtmcdchpignfsfsqcafcscgntlgieetccpgtgnwsspccqv 255  
 |||  
 OY 61 IQCEPLAPDLGIMNCSPHPLASFSTACTGCTGCTGLKKTTCSSGSGWSNPSPIC 120  
 |||  
 DB 256 iqceplapdlgimncshplastfctactlcegtellgkktlccssgswnpapic 315  
 |||  
 OY 121 QKDKSSSMIKE 132  
 |||  
 DB 316 qkdkssismike 327  
 |||

RESULT 27  
 AAR98121  
 ID AAR98121 standard; Protein; 371 AA.  
 AC AAR98121;  
 DT 01-NOV-1996 (first entry)  
 DE Human lymphocyte cell surface glycoprotein (HULHR) variant.  
 KW Immunoglobulin; transmembrane receptor; adhesion; targeting;  
 KW diagnosis; therapy; drug delivery; antiviral; neuromodulator;  
 KW immunomodulator; cell adhesion; graft rejection; inflammation;  
 KW metastasis.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Region 1..37  
 FT Domain /label= Signal region.  
 FT Domain 38..154  
 FT Domain /label= Lectin domain.  
 FT Domain 159..192  
 FT Binding-site /label= EGF domain.  
 FT Binding-site 196..257  
 FT Binding-site /label= Complement binding repeat 1.  
 FT Binding-site 258..316  
 FT Binding-site /label= Complement binding repeat 2.  
 FT Domain 332..354







FT		/label= Transmembrane domain.
FT	Domain	355..371
XX		/label= Cytoplasmic domain.
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PE	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808122.
PR	08-DEC-1991;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GETH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA:	
DR	WPI; 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
FT	diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure; Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents; or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AAR98106 are given in	
CC	AAR98109-R98135. This variant contains a Phe24Met substitution.	
XX		
SQ	Sequence 371 AA;	
	Query Match	99.2%; Score 727; DB 17; Length 371;
	Best Local Similarity	99.2%; Pred. No. 5.5e-63;
	Matches 131; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
OY	1 CEPLEAPELGTDCNHPRENFSSFOCAFSCSEGNLTGIIEETGGPFCFNWSSPEPTCOV	60
DB	196 cepleapelpgundchtpfnfistssqcatcscegnltlyieetcgpmnwsspeptcyv	255
OY	61 IQCEPLSAPDLTINMNCSHPLASFSTSACTFICSEGTETLGKRTTCESSGIWSNPSPYC	120
DB	256 lqceplsapdldgimncshplastfsactfcsegtelgkrttcessgikwspspic	315
OY	121 OKLDKSFMSIKE 132	
DB	316 qklDKsfsmike 327	
RESULT 31		
ID	AAR98123	
XX	AAR98123 standard; Protein; 371 AA.	

AC	AA98123.	
XX	01-NOV-1996	(first entry)
XX		
XX	Human lymphocyte cell surface glycoprotein (HULHR) variant.	
XX		
KW	Immunoglobulin: transmembrane receptor; adhesion; targeting;	
KW	diagnosis; therapy; drug delivery; antiviral; neuromodulator;	
KW	immunomodulator; cell adhesion; graft rejection; inflammation;	
KW	metastasis.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Region	1..37
FT		/label= Signal region.
FT	Domain	38..154
FT		/label= Lectin domain.
FT	Domain	159..192
FT		/label= EGF domain.
FT	Binding-site	196..257
FT		/label= Complement binding repeat 1.
FT	Binding-site	258..316
FT		/label= Complement binding repeat 2.
FT	Domain	332..354
FT		/label= Transmembrane domain.
FT	Domain	355..371
FT		/label= Cytoplasmic domain.
XX		
PN	US5514582-A.	
XX		
PD	07-MAY-1996.	
XX		
PE	23-FEB-1989;	89US-0315015.
XX		
PR	22-NOV-1989;	89US-0440625.
PR	23-FEB-1989;	89US-0315015.
PR	16-DEC-1991;	91US-0808121.
PR	08-DEC-1992;	92US-0986931.
PR	21-JAN-1994;	94US-0185670.
XX		
PA	(GEPH ) GENENTECH INC.	
XX		
PI	Capon DJ, Lasky LA;	
XX		
DR	WPI: 1996-238773/24.	
XX		
PT	Nucleic acid encoding hybrid immunoglobulin comprising the ligand	
PT	binding site of a receptor fused to Ig constant region - useful for	
PT	diagnosis and treatment e.g. of inflammation	
XX		
PS	Disclosure: Page 19; 41pp; English.	
XX		
CC	A hybrid immunoglobulin chain comprising the ligand binding site of	
CC	a single transmembrane receptor without an active transmembrane	
CC	region; fused at its C-terminus with the N-terminus of an	
CC	immunoglobulin constant region. The receptor is not a member of the	
CC	immunoglobulin super family, nor a multiple subunit polypeptide	
CC	encoded by discrete genes. The hybrid immunoglobulin chain combines	
CC	the adhesion/targeting of a ligand binding partner (LBP) with the	
CC	effector functions of immunoglobulin and can bind to and/or activate	
CC	more than one ligand. It can be used diagnostically for the in	
CC	vitro assay of LBP and their targets; or therapeutically to deliver	
CC	LBP such as toxins, enzymes, growth factors to particular cells.	
CC	Typical applications are as antiviral, neuromodulating and	
CC	immunomodulating agents, or as modulators of cell adhesion (e.g. in	
CC	treating graft rejection; inflammation; metastasis of lymphoma etc.)	
CC	The immunoglobulin component increases plasma half life and	
CC	facilitates purification while deletion of the transmembrane region	
CC	facilitates recovery, improves aqueous solubility and removes	
CC	potentially immunogenic epitopes. Variants of the human lymphocyte	
CC	cell surface glycoprotein described in AA98106 are given in	
CC	AA98109-898135. This variant contains a Lys298-Lys299; Arg-Arg	







XX WP1: 1992-056864/07.  
DR N-PSDB: AAQ22500.  
XX  
XX  
PT New CD53 cell surface antigen and DNA encoding it - for  
PT immuno-therapy and diagnosis of haematopoietic neoplasms, etc.  
XX  
XX  
PS Example 14; Page 106; 160pp; English.  
XX  
XX Two cDNA clones encoding Leu8 determinants were isolated from a  
CC human T cell library using the rapid immunoselection cloning method  
CC (see e.g. AAQ21164 for description of method). This protein sequence  
CC (see e.g. AAQ21164 for description of method). This protein sequence  
CC C-terminal domain is characteristic of surface proteins that are  
CC attached to the cell membrane by covalent linkage to a  
CC phosphatidylinositol-substituted glycan.  
CC See AAQ21184 for the larger insert and AAR20815 for the major form of  
CC the Leu8 antigen that it encodes.  
XX  
XX  
SQ Sequence 363 AA;

Query Match	Similarity	98.13:	Score 719:	DB 13:	Length 363:
Best Local	Similarity	98.58:	Pred. No. 3.2e-62:		
Matches 130:	Conservative	0:	Mismatches 2:	Indels 0:	Gaps
Oy	1	CEPLAPAPLGMDCTHHPGNFSSSOCAFCSCSECTNLGIEETTCGPFGNMSSPEPTCOV	60		
Db	210	cepleapeligmdcthsignstfsqcafsesegntltglettcgpfgnmsspeptcyv	269		
Oy	61	IQCEPLSAPDGLIMNCSHPLASFSTSACTFFICSEGTELIGKKTICSSGIMNSPPTC	120		
Db	270	iqceplsapdglimncshplasfstaactficsegtelgkktlcassgimnsppic	329		
Oy	121	QKDKSFSMIKE	132		
Db	330	qkdksfsmike	341		
RESULT	37				
AA091443	ID	AA091443	standard; Protein: 363	AA.	
AA	AA091443:				
XX	31-OCT-1996	(first entry)			
XX	Human Leu8 antigen.				
XX	Cell surface antigen; cloning; immunoselection; immunotherapy;				
KW	therapy; diagnosis; vector; Leu8 antigen; T-Lymphocyte;				
KW	antiinflammatory.				
XX	Homo sapiens.				
XX	Key	Location/Qualifiers			
FT	Modified-site	113..115			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	156..158			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	229..231			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	269..271			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	285..287			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	299..301			
FT	Modified-site	/label= Glycosylation_site			
FT	Modified-site	324..326			
FT	Modified-site	/label= Glycosylation_site			
XX	US5506126-A.				
XX					

PD	09-APR-1996.	
XX		
PF	25-FEB-1988;	88US-0160416.
XX		
PR	01-DEC-1992;	92US-0983647.
PR	25-FEB-1988;	88US-0160416.
PR	13-JUL-1989;	89US-0379076.
PR	13-JUL-1990;	90US-0553759.
PR	18-OCT-1993;	93US-0139273.
XX		
PA	(GEHO ) GEN HOSPITAL CORP.	
XX		
PI	Aruffo A, Seed B;	
XX		
DR	WPI, 1996-200279/20.	
DR	N-PSDB; AAT14723.	
XX		
PT	Cloning of cDNA encoding cell surface antigen - useful for isolation	
PT	of diagnostic and therapeutic proteins	
XX		
PS	Example 14; Column 71-74; 79pp; English.	

CC	2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were isolated from a human T-cell library using a novel immunoselection cloning method. The longer insert (AAR14723) contained 2,350 residues, while the shorter lacked 436 internal residues. A major transcript of 2.4 kb was present in peripheral blood mononuclear cells (PBMC), tonsillar B cells, and several lymphocytic cell lines, and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukaemic T-cell lines. Leu8 antigens can be obtained for diagnostic and therapeutic use. The presence or absence of Leu8 on CD4+ T-cells identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets
CC	Soluble forms of Leu8 can act as antiinflammatory agents by reducing lymphocyte migration.
CC	
XX	
XX	
SQ	Sequence 363 AA:
Query Match	98.1%; Score 719; DB 17; Length 363;
Best Local Similarity	98.5%; Pred. No. 3.2e-62;
Matches 130; Conservative	0; Mismatches 2; Indels 0; Gaps
OY	1 CEPLEAPELGTMDCTHPRGNGFNSFSQCAFSCSEGTNLGIEETTCGPGNNSSPEPTCOV 60       
Dd	210 cepleapelgtmcdcthsignfsfscqafscsegnltgiettcogpgnwspeptcoy 260 
OY	61 IOCEPLSAPADLGIMNCSHPLASFSFTSACTPFCSGTGLIGKKKTIICSSSIWNSPSTIC 120       
Dd	270 iqceplsapadlgimncshplasfstactfcsegeltlqkkklicessgiwnspstic 320 
OY	121 OKLDKSFMIKE 132 
Dd	330 qldksfsmike 341 
RESULT 38	
AAY96184	
ID	AAY96184 standard; Protein: 363 AA.
XX	
AC	AAY96184;
XX	
DT	19-DEC-2000 (first entry)
XX	
DE	Human T-cell specific Leu8 antigen.
XX	
KW	Leu8; cell surface antigen; human; immunoselection; panning; immunodiagnosis; diagnosis; immunotherapy; gene therapy; immune disorder; infection; asthma; immune-complex disease; amyloidosis; multiple sclerosis; inflammation; antiinflammatory.
XX	
OS	Homo sapiens.
PN	US611093-A.
XX	





```

QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGIRETTCGPRGNMSSPEPTQV 60
DB 210 cepleaprlgmdcthsignfsfsqcfscsegtnlgiectcpgfnwsspeptqv 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFTCSGTELGKKTICSSGIMSNPSPIC 120
DB 270 lqceplaspdlgimncshplasfstsactftcsegtelgkkticssgimwnpplc 329
QY 121 QKLDKSFMSIKE 132
DB 330 qkldksfsmike 341

RESULT 40
AAR91442
ID AAR91442 standard; Protein: 385 AA.
XX
AC AAR91442;
XX
DT 31-OCT-1996 (first entry)
XX
DE Human Leu8 antigen.
XX
KW Cell surface antigen; cloning: immunoselection; immunotherapy;
KW therapy; diagnosis; vector; Leu8 antigen; T-lymphocyte;
KW antinflammatory.
XX
OS Homo sapiens.
XX
FH Key location/Qualifiers
FT Modified-site 113..115
FT Modified-site /label= glycosylation_site
FT Modified-site 136..158
FT Modified-site /label= glycosylation_site
FT Modified-site 229..231
FT Modified-site /label= glycosylation_site
FT Modified-site 269..271
FT Modified-site /label= glycosylation_site
FT Modified-site 285..287
FT Modified-site /label= glycosylation_site
FT Modified-site 299..301
FT Modified-site /label= glycosylation_site
FT Modified-site 324..326
FT Modified-site /label= glycosylation_site
FT Domain 347..368
FT /label= Transmembrane_domain
XX
PN US5506126-A.
XX
PD 09-APR-1996.
XX
PF 25-FEB-1988; 88US-0160416.
XX
PR 01-DEC-1992; 92US-0983647.
PR 25-FEB-1988; 88US-0160416.
PR 13-JUL-1989; 89US-0379076.
PR 13-JUL-1990; 90US-0553759.
PR 18-OCT-1993; 93US-0139273.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Arufo A., Seed B;
XX
DR WPI: 1996-200279/20.
DR N-PSDB; AAT14723.
XX
PT Cloning of cDNA encoding cell surface antigen - useful for isolation
PT of diagnostic and therapeutic proteins
XX
PS Example 14; Column 71-74; 79pp; English.
XX
CC 2 cDNA clones encoding Leu8 determinants (AAR91442 and AAR91443) were

```

```

CC isolated from a human T-cell library using a novel immunoselection
CC cloning method. The longer insert (AAT14723) contained 2,350
CC residues, while the shorter lacked 436 internal residues. A major
CC transcript of 2.4 kb was present in peripheral blood mononuclear
CC cells (PBMC), tonsillar B cells, and several lymphocytic cell lines,
CC and a minor transcript of 2.0 kb in PBMC, Jurkat and HSB-2 leukemia
CC T-cell lines. Leu8 antigens can be obt. for diagnostic and
CC therapeutic use. The presence or absence of Leu8 on CD4+ T-cells
CC identifies suppressor-inducer and helper-inducer CD4+ T-cell subsets.
CC Soluble forms of Leu8 can act as antinflammatory agents by reducing
CC lymphocyte migration.
XX
SO Sequence 385 AA;

```

```

Query Match 98.1%; Score 719; DB 17; Length 385;
Best Local Similarity 98.5%; Pred. No. 3,56-62;
Matches 130; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 CEPLAPRLGMDCTHPRGNFSFSQCAFSCSEGTNLGIRETTCGPRGNMSSPEPTQV 60
DB 210 cepleaprlgmdcthsignfsfsqcfscsegtnlgiectcpgfnwsspeptqv 269
QY 61 IQCEPLAPDLGIMNCSHPLASFSTSACTFTCSGTELGKKTICSSGIMSNPSPIC 120
DB 270 lqceplaspdlgimncshplasfstsactftcsegtelgkkticssgimwnpplc 329
QY 121 QKLDKSFMSIKE 132
DB 330 qkldksfsmike 341

```

```

Search completed: September 7, 2002, 10:14:09
Job time: 267 sec

```

**This Page Blank (uspto)**

**This Page Blank (uspto)**